Dispositivos activos y componentes pasivos para sistemas DWDM

Carlos Janer Jiménez
Departamento de Ingeniería Electrónica
Escuela Superior de Ingenieros de Sevilla
Universidad de Sevilla
email: janer@zipi.us.es;janer@us.es

Versión provisional

Si quieren expresar comentarios, hacer sugerencias o comunicar errores y/o erratas, contacten, por favor, con el autor.

Todas las sugerencias serán bienvenidas.

Junio 2009

Preámbulo

Esta es la primera versión de un libro que, en otras circunstancias, habría tenido otro título. Mi primera intención fue escribir conjuntamente con el Dr. D. Alejandro Carballar Rincón un texto sobre *componentes* y *sistemas* DWDM que, de alguna manera, reflejara la docencia que hasta hace un año y, durante al menos quince, estuve impartiendo en la asignatura de Comunicaciones Ópticas de la titulación de Ingeniero de Telecomunicación de la Escuela de Ingenieros de Sevilla. Sin embargo las numerosas obligaciones de mi colega le han impedido redactar, como era su deseo, la parte correspondiente a sistemas. Quiero aprovechar la ocasión para expresarle mi profundo agradecimiento ya que el contenido de los capítulos doce, trece, catorce y quince se ha basado en unos apuntes manuscritos suyos y, además, ha desarrollado los programas que han permitido hacer todas las simulaciones numéricas que aparecen en el libro.

No me siento satisfecho del resultado final y no es sólo por las numerosos errores y erratas que a buen seguro contiene ni por su pobre redacción. Mi frustración está provocada por los temas que **no** he podido incluir, entre los que destaco, sin ser exhaustivo, los convertidores en longitud de onda, los moduladores ópticos por electroabsorción, las comunicaciones ópticas coherentes y todos los dispositivos basados en semiconductores de bajas dimesiones (pozos, líneas y puntos cuánticos). Creo que la mayor parte de los alumnos de un quinto curso de Igeniero de Telecomunicación carecen de los conocimientos físicos necesarios para abordar estas cuestiones. Asímismo debo advertir que los dispositivos que sí se han estudiado se han modelado de una manera muy burda y que lo único que se pretende con estos modelos tan simplistas es permitir es que se entiendan sus principios de funcionamiento.

Trataré de seguir puliendo el texto y subsanando errores a medida que se me vayan comunicando. Agradezco de antemano a los lectores su colaboración en la mejora del contenido de este libro.

Por último agradezco a mi mujer, Eli, el mero hecho de que exista y quiera compartir su vida conmigo.

Carlos Janer.
Sevilla, Junio del 2009.

Índice general

Capítulo 1

Propagación de pulsos por medios dispersivos

Resumen: Antes de estudiar la propagación de pulsos luminosos por una fibra óptica, debemos entender primero cómo los fenómenos de atenuación y dispersión degradan y distorsionan las señales que se propagan por un medio dieléctrico. Estos procesos físicos aparecen incluso si sus dimensiones transversales son infinitas, es decir, en medios en los que la luz no está confinada. El origen último de la atenuación y la dispersión, reside en la interacción de la luz con las partículas cargadas (electrones y núcleo) de los átomos del medio. Estos átomos los vamos a modelar mediante un conjunto de osciladores armónicos (modelo de Lorentz) en los que el campo eléctrico ejerce una fuerza que tiende a separar estas partículas cargadas e induce, por tanto, dipolos microscópicos. Los dipolos absorben y emiten radiación electromagnética produciéndose, así, la atenuación y dispersión de los pulsos luminosos que se propagan por el medio material.

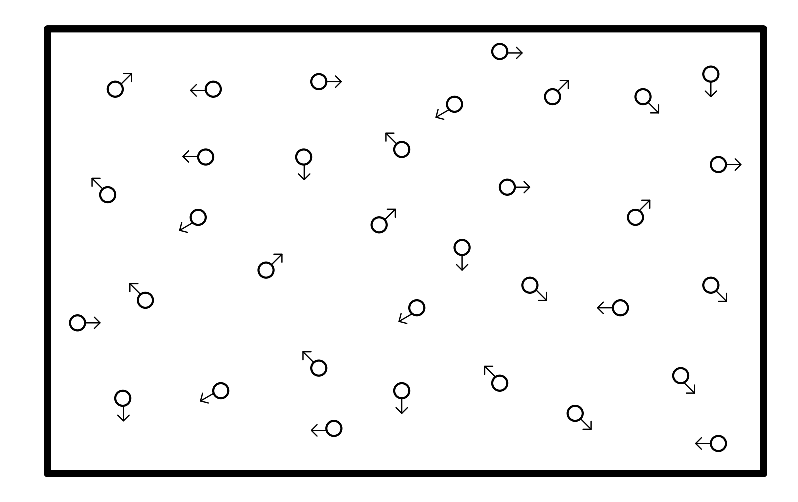


Figura 1.1: Átomos de una gas atómico en un recinto cerrado

1.1. Introducción

En este capítulo vamos a estudiar la propagación de pulsos luminosos por medios dieléctricos de dimensiones transversales infinitas [1]. Como consecuencia de la interacción entre la luz y el medio dieléctrico aparecen dos fenómenos físicos que son:

1. La absorción.

2. La dispersión.

Si el medio de propagación fuese el vacío, la atenuación y la dispersión no se producirían. Estos fenómenos surgen, en última instancia, por la interacción entre la radiación luminosa y las partículas cargadas que constituyen el medio material.

Para abordar este estudio, vamos a hacer una serie de hipótesis simplificativas:

1. Vamos a estudiar, en un principio, la propagación de ondas planas (para no tener que estudiar fenómenos difractivos) perfectamente monocromáticas. Estas ondas constituyen una base de descomposición en modos propios espaciales (la más adecuada para un medio de dimensiones infinitas) sobre la que proyectar el pulso luminoso cuya propagación queremos estudiar. Además, puesto que por el canal (o medio de propagación) se pretende enviar información en una determinada dirección (por ejemplo el eje z), resulta conveniente definir un sistema de ejes cartesianos en el que las superficies equifase sean perpendiculares a esta dirección.

2. La portadora de la información es perfectamente monocromática y se va a modular en amplitud por medio de una envolvente de duración finita. Por tanto el ancho de banda del pulso está fijado por la anchura espectral de la transformada de Fourier de la envolvente de dicho pulso.

[1]Los fenómenos de guiado de la luz, característicos de medios de dimensiones transversales finitas (tales como la fibra óptica) se abordan en temas posteriores.

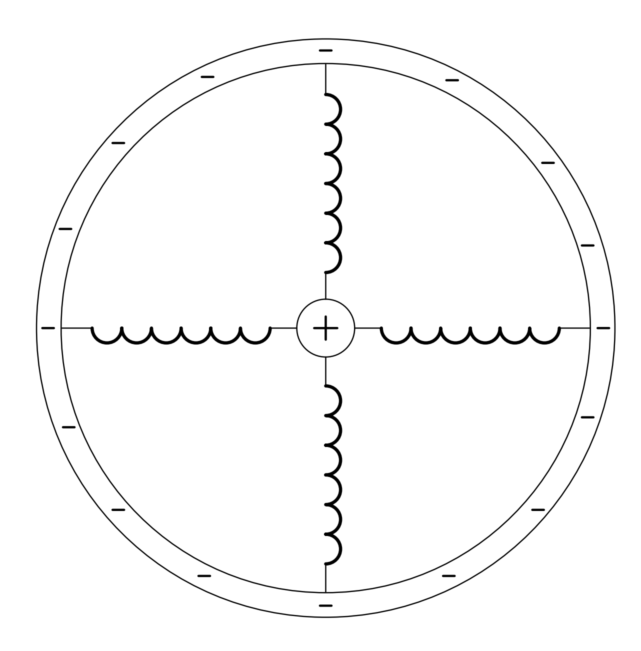


Figura 1.2: Modelo mecánico tridimensional del átomo

Vemos, entonces, que para abordar el estudio de la propagación de pulsos luminosos, es necesario analizar previamente la propagación de ondas totalmente monocromáticas, ya que éstas constituyen su base de descomposición natural. Suponemos que el medio de propagación tiene una respuesta lineal y, por tanto, que podemos utilizar la técnica de la transformada de Fourier para analizar la propagación del pulso. En primer lugar, expresaremos este pulso como una suma infinita (integral de Fourier) de ondas planas monocromáticas. A continuación propagaremos independientemente cada una de estas componentes para, por último, sumar (integrar) cada una de éstas y obtener, así, el pulso al otro extremo del canal (medio de propagación). En términos matemáticos, el pulso propagado se expresa como la antitransformada de Fourier del producto de la función de transferencia del canal (que describe como se propagan las componentes monocromátricas) multiplicada por la transformada de Fourier del pulso a la entrada. Este método es aplicable sólo si la interacción entre el campo electromagnético y el medio material se puede describir por medio de un conjunto de ecuaciones diferenciales lineales. Hacia el final de este libro, relajaremos la hipótesis de linealidad del medio de propagación.

1.2. Modelo del medio dieléctrico

El sílice es un sólido vítreo, cuya estructura microscópica estudiaremos en un capítulo posterior [2]. Para simplificar el análisis, vamos a considerar que el medio dieléctrico material es un gas atómico diluido. Los fenómenos de absorción y dispersión que deduciremos de este estudio, serán cualitativamente parecidos a los que se producen en el sílice.

En la fig. {1.1} se muestra un recinto cerrado que contiene átomos de un gas a

[2]Veremos que por tratarse de un material amorfo, aparece un fenómeno físico adicional que provoca atenuación: el scattering Rayleigh.

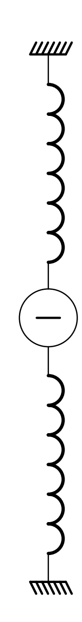


Figura 1.3: Modelo mecánico unidimensional del átomo

baja presión. Los procesos de interacción luz–materia se producen porque los átomos están constituidos por partículas cargadas capaces de generar dipolos eléctricos en presencia de la luz (el campo electromagnético ejerce fuerzas sobre las partículas cargadas y se produce una separación entre las cargas positivas y negativas, induciéndose, de esta manera, dipolos microscópicos). Estos dipolos microscópicos, convenientemente sumados (es decir, teniendo en cuenta las colisiones que se producen entre los átomos), dan lugar a un momento dipolar macroscópico a partir del cual podemos calcular el coeficiente de atenuación y el índice de refracción. Estas dos magnitudes, que dependen de la pulsación de la onda que se propaga por el medio, permiten calcular la pérdida de energía que se produce durante la propagación (la atenuación) y el ensanchamiento temporal de los pulsos (dispersión).

En la fig. {1.2} se muestra el modelo mecánico del átomo (modelo de Lorentz) [11], en el que la región central representa el núcleo del átomo (donde se encuentra concentrada la carga positiva). Los muelles representan las fuerzas electromagnéticas que mantienen a los electrones (cargas negativas) orbitando alrededor del núcleo. Vamos a considerar que la luz incidente está linealmente polarizada [3], por lo que el campo eléctrico vibra en una determinada dirección del espacio, dirección en la que se va a producir la separación entre cargas positivas y negativas y la aparición de un momento dipolar microscópico. Por este motivo podemos emplear el modelo unidimensional mecánico del átomo [11] mostrado en la fig. {1.3}. La ecuación de movimiento de la carga negativa (consideramos que la carga positiva está fija por ser mucho más masiva) es la siguiente:

$$\sum_i F_i = m\ddot{x} \Rightarrow m\ddot{x} = -kx - qE(t) \Rightarrow \ddot{x} + \frac{k}{m}x = -\frac{q}{m}E(t) \tag{1.1}$$

[3] En realidad basta con que el tiempo de coherencia de la luz sea grande en comparación con el tiempo característico de respuesta de los dipolos inducidos.

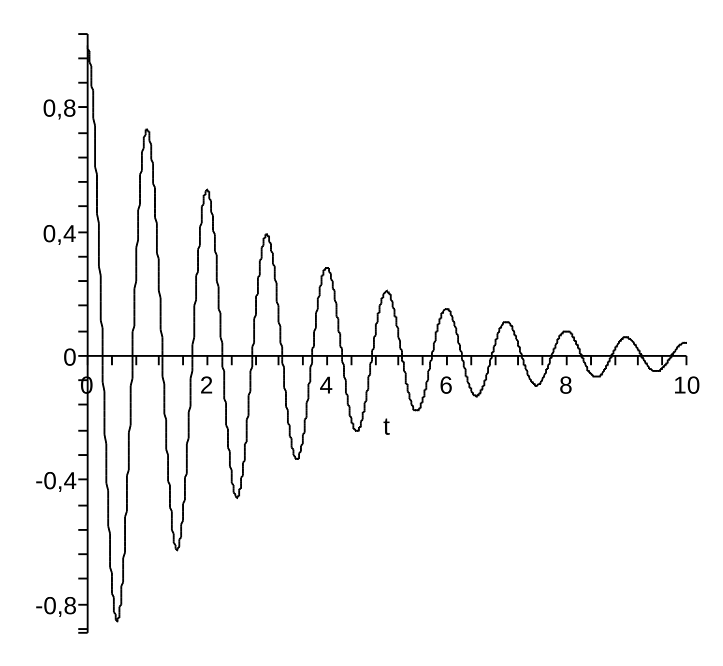


Figura 1.4: Evolución temporal del oscilador. Representamos $\frac{x(t)}{x_\circ}$ en el eje vertical frente a $\frac{2\pi t}{T}t$ en el eje horizontal.

donde $E(t)$ es el campo eléctrico, k es la constante de elongación del muelle, q y m son la carga y la masa del electrón. Ahora bien, esta ecuación no tiene en cuenta la interacción que se produce entre el dipolo inducido en el átomo y el campo electromagnético creado por él mismo. Este fenómeno de autointeracción se puede estudiar de una manera relativamente sencilla aplicando el principio de conservación de la energía. Aunque el razonamiento no es complicado, no lo vamos a reproducir aquí y remitimos al lector interesado a la bibliografía [11]. La conclusión de ese estudio es que aparece un nuevo término en la ecuación diferencial que es proporcional a la velocidad y que representa, por tanto, una disipación de energía.

$$\ddot{x} + \frac{\Gamma}{m}\dot{x} + \frac{k}{m}x = -\frac{q}{m}E(t) \tag{1.2}$$

donde Γ es el coeficiente de fricción del muelle. Si ahora definimos la pulsación característica como $\omega_\circ = \sqrt{\frac{k}{m}}$ y el coeficiente de fricción normalizado como $\gamma = \frac{\Gamma}{m}$ nos queda finalmente la siguiente ecuación diferencial:

$$\ddot{x} + \gamma\dot{x} + \omega_\circ^2 x = -\frac{qE(t)}{m} \tag{1.3}$$

Esta ecuación diferencial es sobradamente conocida en la matemática física. Describe la *vibración forzada* de un sistema mecánico, cuya pulsación propia de oscilación es $\omega_\circ$ y cuyo coeficiente de disipación es γ. Si el campo electromagnético incidente $E(t)$ tiene una

pulsación próxima a $\omega_\circ$, el sistema entra en *resonancia* con este estímulo externo, roba energía electromagnética y la emplea en aumentar su energía mecánica. Comprobémoslo [4]:

La relación $\ddot{x}+\gamma\dot{x}+\omega_\circ^2 x = 0$ es la ecuación homogénea asociada a (1.3), su polinomio característico es $\xi^2+\gamma\xi+\omega_\circ^2 = 0$ y las raíces de este polinomio son: $\xi = \frac{-\gamma\pm\sqrt{\gamma^2-4\omega_\circ^2}}{2} \simeq -\frac{\gamma}{2}\pm j\omega_\circ$ donde hemos tenido en cuenta que el coeficiente de fricción es pequeño. Por tanto, la solución de la ecuación homogénea asociada a (1.3) es (tomando una referencia de fase adecuada):

$$x(t) = X_o e^{-\frac{\gamma}{2}t}\cos(\omega_\circ t) \tag{1.4}$$

Vemos que el dipolo tiene un modo propio de oscilación de pulsación $\omega_\circ$. Esta vibración está amortiguado por la interacción entre el campo electromagnético creado por el dipolo consigo mismo, por lo que el sistema va perdiendo energía mecánica progresivamente, tal como se muestra en fig. {1.4}. Si $U(t)$ es la energía mecánica total del oscilador, se cumple, entonces, la siguiente ecuación:

$$U(t) = U(t)^{pot} + U(t)^{cin} = \frac{1}{2}kx(t)^2 + \frac{1}{2}m\dot{x}^2 = U(t_\circ)e^{-\gamma t} \tag{1.5}$$

¿Qué ocurre si tenemos un campo electromagntético incidente que oscila a una pulsación cercana a $\omega_\circ$? Pues que el átomo entra en resonancia, produciéndose una vibración *forzada* que roba rápidamente energía a la luz incidente, aumentando su energía mecánica. Este proceso físico se conoce con el nombre de *absorción*. La energía mecánica perdida por el término de fricción del dipolo se traduce en la emisión de fotones de pulsación cercana, pero no necesariamente igual, a $\omega_\circ$ y que se propagan en cualquier dirección, con cualquier estado de polarización y cualquier referencia de fase. Este proceso físico se conoce con el nombre de *emisión espontánea*. En realidad el dipolo inducido puede perder energía por otros procesos físicos que no implican la emisión de fotones [15]. El estudio microscópico de los fenómenos de desexcitación, aunque resulta muy interesante, se sale fuera de los límites de este libro.

En la fig. {1.5} vemos cómo la luz incidente, de pulsación $\omega \simeq \omega_o$, se atenúa y cómo se generan fotones por emisión espontánea. Estos fotones tiene características distintas de los absorbidos, ya que el estado de polarización, dirección de propagación, referencia de fase y, hasta cierto punto, pulsación, son aleatorios.

1.3. Momentos dipolar microscópico y macroscópico

Para describir cuantitativamente los procesos de interacción entre la luz y el medio atómico, tenemos que calcular el momento dipolar macroscópico sumando los momentos

[4] Por supuesto que esta ecuación diferencial se puede integrar exactamente empleando el método de la transformada de Fourier. De hecho esto es lo que haremos con una ecuación análoga a ésta un poco más adelante.

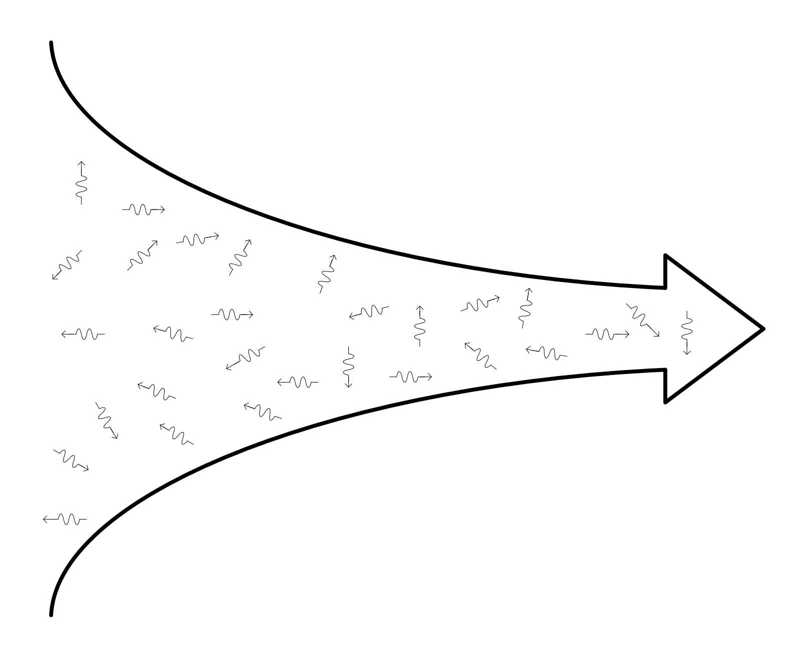


Figura 1.5: Atenuación y emisión espontánea de luz

dipolares microscópicos producidos por los átomos individuales. A la hora de calcular esta suma, debemos tener en cuenta que los átomos están chocando continuamente entre sí, ya que se encuentran sometidos a un movimiento de agitación térmica (la temperatura ambiente es del orden de 300K). Estas colisiones producen saltos de fase aleatorios [92] en la radiación emitida por cada átomo (la fase instantánea después de una colisión está incorrelada con respecto a la que tenía antes del choque). Por este motivo no podemos sumar directamente (como si todas estuvieran en fase) las contribuciones de los átomos individuales [5].

El objetivo de este apartado es, precisamente, calcular el vector polarización macroscópico a partir de los momentos dipolares microscópicos inducidos en los átomos. Determinemos, primero, el momento dipolar de un átomo aislado:

$$\left.\begin{array}{r}\ddot{x}+\gamma\dot{x}+\omega_\circ^2 x=-\dfrac{qE(t)}{m}\\ \mu_x(t)=-qx(t)\end{array}\right\}\Rightarrow\left.\begin{array}{r}\ddot{\mu_x}+\gamma\dot{\mu_x}+\omega_\circ^2\gamma_x=\dfrac{q^2E(t)}{m}\\ t=0;\ \mu_x(t)=\mu_{x\circ},\ E_x=0\end{array}\right\}\Rightarrow$$

$$\mu_x(t) = \mu_{x\circ}e^{-\frac{\gamma}{2}t}\cos(\omega_\circ t+\Phi_\circ) \qquad (1.6)$$

Si todos los osciladores estuviesen inicialmente en fase y no hubiese colisiones entre átomos, el cálculo del momento dipolar macroscópico sería muy sencillo, sólo habría que

[5] Para tiempos cortos con respecto a las constantes de relajación del sistema, se produce las llamadas oscilaciones de Rabi [11], en las que se produce un intercambio periódico de energía entre el campo electromagnético y el medio material.

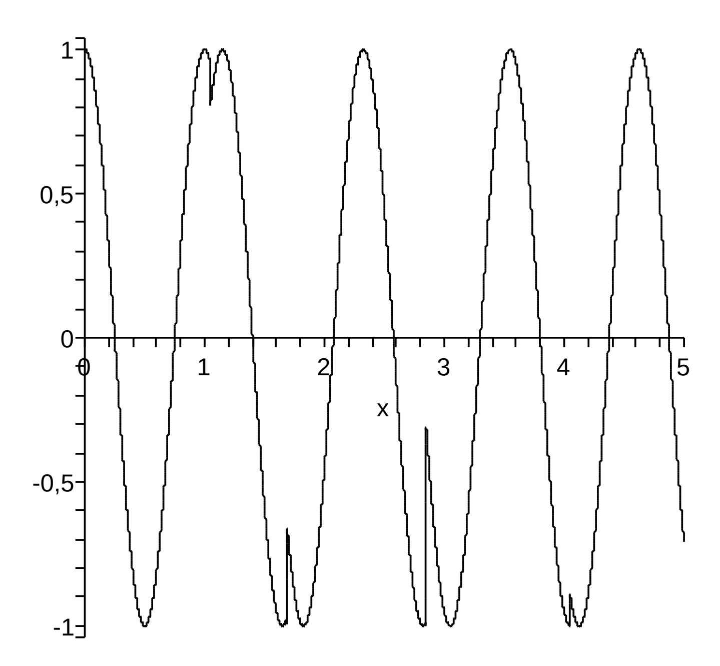


Figura 1.6: Saltos de fase en el momento dipolar microscópico producidos por las colisiones. En el eje vertical se representa el momento dipolar microscópico normalizado $\frac{\mu_x}{\mu_{x\circ}}$ y en el horizontal el tiempo normalizado $\frac{2\pi t}{T}$.

multiplicar por la densidad de átomos, $P(t) = \frac{N\mu_x(t)}{dV}$. Como hay colisiones (que vamos a suponer elásticas), tenemos que la evolución temporal del momento dipolar microscópico de un átomo aislado es, aproximadamente, la descrita por la fig. {1.6}.

El momento dipolar macroscópico decae mucho más rápidamente que el microscópico. Esto es debido a que tras una colisión la fase del dipolo se hace aleatoria, distribuyéndose uniformemente en el intervalo $(0, 2\pi)$. Los campos electromagnéticos emitidos por los dipolos microscópicos que, en un principio interferían constructivamente, comienzan a interferir destructivamente. Se puede probar fácilmente que si T es la variable aleatoria que describe el tiempo que transcurre entre dos colisiones sucesivas de un dipolo (átomo) y $T_\circ$ es la media de esta variable aleatoria, entonces el momento dipolar macroscópico viene descrito por la siguiente ecuación diferencial [92]:

$$\begin{gathered}\ddot{P}_x + (\gamma + \tfrac{2}{T_\circ})\dot{P}_x + \omega_\circ^2 P_x = \frac{Nq^2E(t)}{dVm} \\ T_\circ := < T > \end{gathered} \tag{1.7}$$

$T_\circ$ es el tiempo que, en promedio, transcurre entre dos colisiones sucesivas en un átomo. En la fig. {1.7} se muestra el efecto que las colisiones tienen sobre el momento dipolar macroscópico. Se han sumado los momentos dipolares de cuatro átomos que inicialmente vibraban en fase y se han introducido saltos de fases aleatorios (que simulan las colisiones) en cada uno de ellos. Se observa que la suma total de estos momentos dipolares decae más

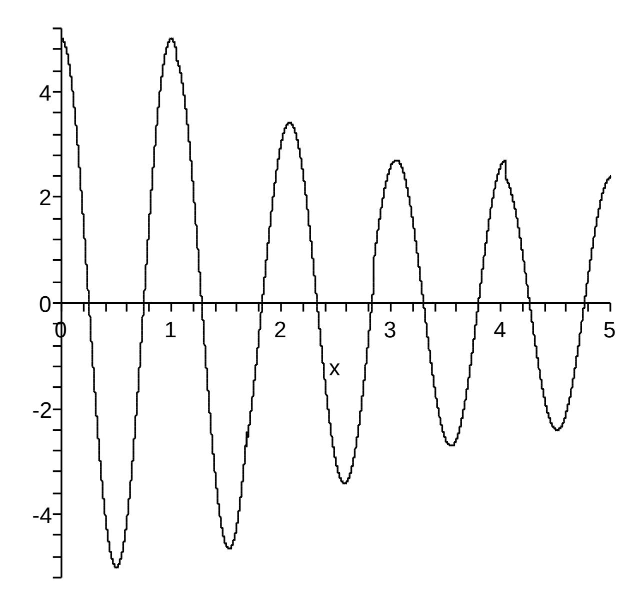


Figura 1.7: Término de decaimiento adicional producido por las colisiones. En el eje vertical se ha representado la suma de los momentos dipolares microscópicos de cuatro átomos que sufren colisiones aleatoriamente, $\frac{1}{\mu_{x\circ}}\sum_{i=1}^{i=4}\mu_{x_i}(t)$ y en el horizontal el tiempo normalizado $\frac{2\pi t}{T}$.

rápidamente que la de un dipolo aislado [6], por lo que las colisiones pueden modelarse introduciendo un término de fricción adicional en la ecuación diferencial del momento dipolar macroscópico, término que no está asociado a una disipación de energía (puesto que las colisiones son elásticas). Conociendo la expresión del momento dipolar macroscópico, estamos en condiciones de analizar los procesos de interacción luz-materia (salvo la emisión estimulada, que la estudiaremos cuando abordemos los principios de funcionamiento de los láseres).

1.4. Susceptibilidad dieléctrica en medios de dos niveles

Aclaremos, en primer lugar, lo que entendemos por un "medio atómico de dos niveles" [11]. Lo que queremos decir es que suponemos que el oscilador atómico tiene una única pulsación propia de oscilación $\omega_\circ$. Esto no es normalmente así [7], pero vamos a suponerlo para facilitar el análisis que sigue.

Supongamos que se propaga una onda electromagnética monocromática de pulsación ω cercana a $\omega_\circ$. El campo vectorial que describe las propiedades eléctricas dentro del medio atómico, es el vector desplazamiento $\vec{D}$ [9] y su expresión es:

[6]Por lo que las oscilaciones de Rabi son observables sólo a escalas temporales muy cortas, a menos que la temperatura del gas sea extremadamente baja.

[7]En realidad, nunca lo es. Siempre hay varias frecuencias de oscilación propias involucradas, ya que los estados energéticos accesibles para los átomos son no sólo dos sino muchos.

$$\vec{D} = \epsilon_\circ \vec{E}(\omega) + \vec{P}(\omega) = \epsilon_\circ \left[1 + \frac{\vec{P}(\omega)}{\epsilon_\circ \vec{E}(\omega)} \right] \vec{E}(\omega) = \epsilon_\circ \epsilon_r(\omega) \vec{E}(\omega)$$
$$\epsilon_r(\omega) = 1 + \chi(\omega), \;\; \chi(\omega) := \frac{\vec{P}(\omega)}{\epsilon_\circ \vec{E}(\omega)} \tag{1.8}$$

$\chi(\omega)$ recibe el nombre de susceptibilidad dieléctrica del medio [9], mientras que $\epsilon_r(\omega)$ es la constante dieléctrica relativa del medio [9]. Ambas son magnitudes complejas, cuyas expresiones podemos determinar a partir de la ecuación diferencial que describe la evolución temporal del momento dipolar macroscópico [8]:

$$\left. \begin{array}{r} \ddot{P}_x + (\gamma + \frac{2}{T_\circ})\dot{P}_x + \omega_\circ^2 P_x = \dfrac{Nq^2 E(t)}{dVm} \\ \tilde{E} := E(t) = E_\circ e^{j\omega t} \end{array} \right\} \;\Rightarrow\; \{-\omega^2 + j(\gamma + \frac{2}{T_\circ})\omega + \omega_\circ^2\}\tilde{P} = \frac{Nq^2 \tilde{E}}{dVm} \Rightarrow$$
$$\chi(\omega) := \frac{\tilde{P}(\omega)}{\tilde{E}(\omega)} \;=\; \frac{Nq^2}{dVm\epsilon_\circ} \frac{1}{(\omega_\circ^2 - \omega^2) + j(\gamma + \frac{2}{T_\circ})} \tag{1.9}$$

Esta resultado debe interpretarse con cierta cautela. Evidentemente, el modelo mecánico de Lorentz del átomo es incorrecto. Para estudiar la susceptibilidad dieléctrica de una manera rigurosa, deberíamos emplear un modelo mecano–cuántico del átomo y, para estudiar su interacción con la radiación electromagnética, tendríamos que cuantificar ésta también. Así entraríamos de lleno en un problema de electrodinámica cuántica [22]–[42]–[50]–[67] que queremos evitar [9].

El hecho de modelar el átomo como un oscilador mecánico y considerar la radiación electromagnética como una onda clásica permite obtener resultados *cualitativamente* válidos, en los que los parámetros $\frac{Nq^2}{dVm\epsilon_\circ}$, $\omega_\circ$ y $\Delta\omega := (\gamma + \frac{2}{T_\circ})$ han de *medirse* experimentalmente.

Si quisiéramos determinar el momento dipolar macroscópico creado por un campo eléctrico dado, sólo tendríamos que calcular la transformada de Fourier de este campo, multiplicarlo por la susceptibilidad dieléctrica y, por último, calcular la antitransformada de Fourier de este producto. Evidentemente, ésto no es más que resolver la ecuación diferencial relativa al momento dipolar macroscópico (una vibración forzada por el término de campo eléctrico).

Vamos a descomponer (1.9) en sus partes real e imaginaria y a aproximar cada una de ellas para valores de pulsación cercanos a la pulsación característica ($\omega \simeq \omega_o$). De esta

[8]Lo que vamos a hacer es tener en cuenta que todos los vectores tienen una dependencia temporal sinusoidal (o, si se quiere, fasorial).

[9]Este tratamiento tiene, sin embargo, una gran ventaja: permite hacer un tratamiento exacto de los procesos de interacción luz–materia, a saber, absorción, emisión estimulada y emisión espontánea y poner de manifiesto la íntima relación que existe entre estos dos últimos procesos.

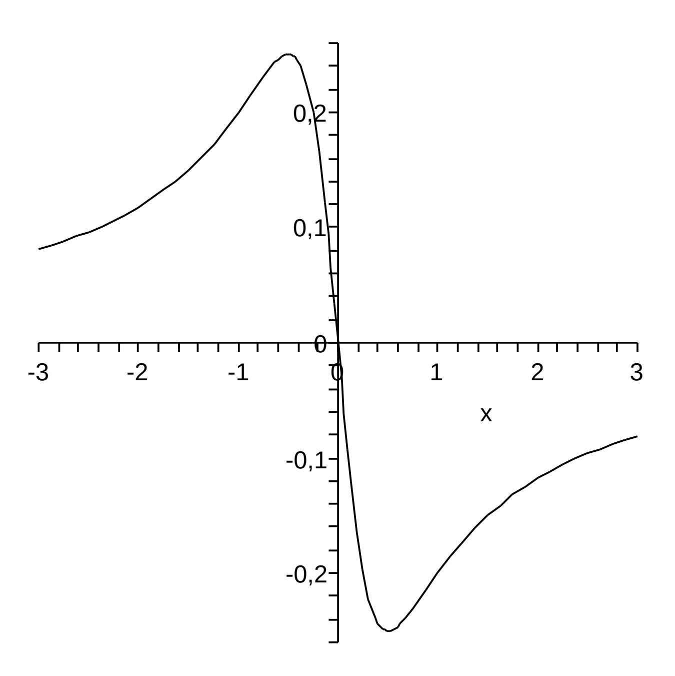


Figura 1.8: Parte real de la susceptibilidad dieléctrica. En el eje vertical se representa la función adimensionalizada $\frac{\mathcal{R}[\chi(\omega)]}{a}$ y en el vertical la pulsación, también adimensionalizada $x = \frac{\omega_\circ - \omega}{\gamma + \frac{2}{T_\circ}}$.

manera tendremos una idea del aspecto que presenta esta función. La expresión resultante depende de una sola variable adimensional y de una única costante:

$$\begin{aligned}
\mathcal{R}\{\chi(\omega)\} &= \frac{Nq^2}{dVm\epsilon_\circ}\frac{\omega_\circ^2 - \omega^2}{(\omega_\circ^2 - \omega^2)^2 + (\gamma + \frac{2}{T_\circ})^2\omega^2} = \\
&= \frac{Nq^2}{dVm\epsilon_\circ}\frac{1}{(\omega_\circ + \omega)(\omega_\circ - \omega) + (\gamma + \frac{2}{T_\circ})^2\frac{\omega^2}{(\omega_\circ+\omega)(\omega_\circ-\omega)}} \simeq \\
&\simeq \frac{Nq^2}{dVm\epsilon_\circ\omega_\circ}\frac{2(\omega_\circ - \omega)}{4(\omega_\circ - \omega)^2 + (\gamma + \frac{2}{T_\circ})^2} \equiv \frac{2ax}{4x^2 + 1} \\
\mathcal{I}\{\chi(\omega)\} &= -\frac{Nq^2}{dVm\epsilon_\circ}\frac{(\gamma + \frac{2}{T_\circ})\omega}{(\omega_\circ^2 - \omega^2)^2 + (\gamma + \frac{2}{T_\circ})^2\omega^2} = \\
&= -\frac{Nq^2}{dVm\epsilon_\circ}\frac{(\gamma + \frac{2}{T_\circ})\omega_\circ}{4\omega_\circ^2(\omega_\circ - \omega)^2 + (\gamma + \frac{2}{T_\circ})^2\omega_\circ^2} \equiv -\frac{a}{4x^2 + 1} \\
x &:= \frac{\omega_\circ - \omega}{\gamma + \frac{2}{T_\circ}} \; a := \frac{Nq^2}{dVm\epsilon_\circ\omega_\circ(\gamma + \frac{2}{T_\circ})} \qquad (1.10)
\end{aligned}$$

En las fig. {1.8} y {1.9} se muestran el aspecto de la parte real e imaginaria de la susceptibilidad dieléctrica. Al hacer la aproximación $\omega \simeq \omega_\circ$ eliminamos el comportamiento

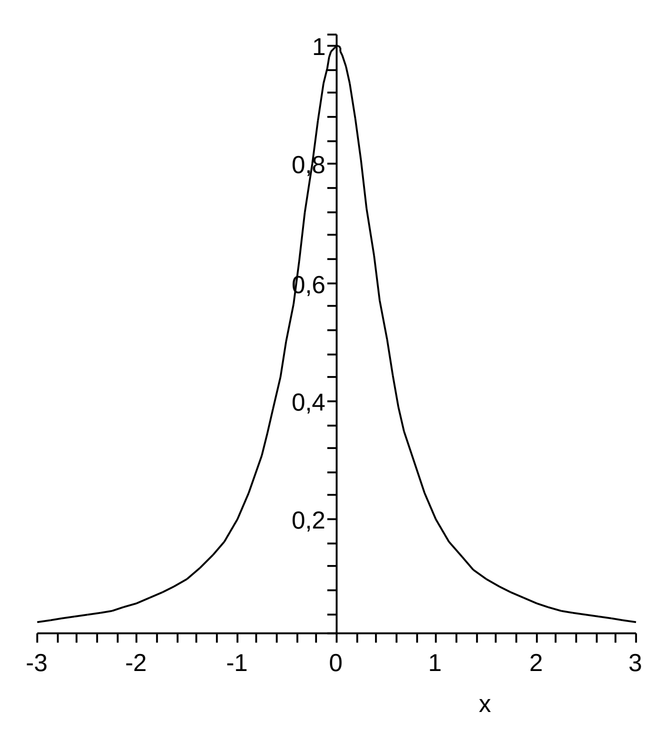


Figura 1.9: Parte imaginaria (cambiada de signo) de la susceptibilidad dieléctrica. En el eje vertical se representa esta función adimensionalizada, $\frac{\mathcal{I}[\chi(\omega)]}{a}$ y en el horinzontal la pulsación, también adimensionalizada $x = \frac{\omega_\circ - \omega}{\gamma + \frac{2}{T_\circ}}$.

asimptótico de la parte real de la susceptibilidad dieléctrica a bajas frecuencias, que tiende a un valor constante que, aproximadamente, es $\frac{Nq^2}{dVm\epsilon_\circ\omega_\circ(\gamma+\frac{2}{T_\circ})}$, mientras que a altas frecuencias tiende a cero. Este resultado es importante porque explica por qué a bajas frecuencias el índice de refracción suele tener valores mayores que a altas frecuencias [49]. A partir de la expresión de la susceptibilidad dieléctrica podemos deducir las características de atenuación y dispersión de un medio atómico. La magnitud física que nos permite determinar la atenuación y dispersión de un paquete de ondas es el *índice de refracción complejo* que es, por definición, igual a la raíz cuadrada de la constante dieléctrica relativa del medio. El comportamiento dispersivo está asociado a la parte real del índice de refracción, mientras que la atenuación está directamente relacionada con la parte imaginaria.

$$\begin{aligned}
\epsilon_r(\omega) &= \epsilon_r^{'}(\omega) + j\epsilon_r^{''}(\omega) = 1 + \mathcal{R}\{\chi(\omega)\} + j\mathcal{I}\{\chi(\omega)\} \\
n(\omega) &:= \sqrt{\epsilon_r^{'}(\omega) + j\epsilon_r^{''}(\omega)} \simeq 1 + \frac{1}{2}\mathcal{R}\{\chi(\omega)\} + j\frac{1}{2}\mathcal{I}\{\chi(\omega)\} \\
n_r(\omega) &\simeq 1 + \frac{1}{2}\mathcal{R}\{\chi(\omega)\},\ n_c(\omega) \simeq \frac{1}{2}\mathcal{I}\{\chi(\omega)\} \qquad (1.11)
\end{aligned}$$

En el razonamiento anterior hemos supuesto que la luz se propaga por un medio muy diluido (por ejemplo un gas a baja presión), de manera que el coeficiente de absorción del medio es bajo, y el índice de refracción cercano a la unidad. Dado que el vector de onda de una onda plana monocomática que se propaga por un medio dieléctrico en la dirección z es

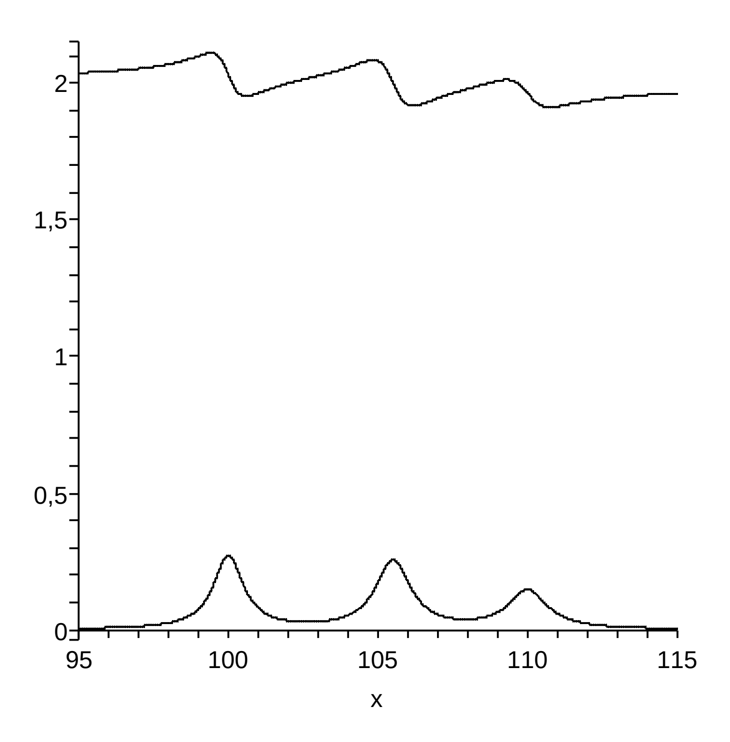


Figura 1.10: Constante dieléctrica relativa en un medio de tres transiciones. Se han representado tanto la parte real como la imaginaria con respecto a la pulsación x dada en Terahertzios.

el producto del vector de onda en el vacío multiplicado por el índice de refracción complejo, llegamos a la conclusión de que su propagación viene descrita por la siguiente expresión:

$$k = k_\circ n(\omega) = k_\circ n_r(\omega) + jk_\circ n_c(\omega) := \beta(\omega) + j\frac{\alpha(\omega)}{2}$$
$$E_z = E_\circ e^{\frac{\alpha(\omega)z}{2}} e^{j[\omega t - \beta(\omega)z]} \tag{1.12}$$

$\alpha(\omega)$ recibe el nombre de coeficiente de atenuación del medio (es un número negativo), mientras que $\beta(\omega)$ es la constante de propagación. La dispersión de los pulsos luminosos (el aumento de su duración temporal a medida que se propagan) se produce, tal como se probará más adelante, porque la constante de propagación, β, varía con repecto a la pulsación, ω, de forma no lineal o, lo que es equivalente, porque el índice de refracción n_r no es constante sino que depende de la pulsación.

1.5. Susceptibilidad dieléctrica en medios de múltiples niveles

Los medios dieléctricos reales tienen múltiples frecuencias de resonancia $\omega_1, \omega_2, ..., \omega_n$, por lo que en realidad la susceptibilidad dieléctrica depende de la pulsación de la siguiente manera [49]:

$$\chi(\omega) = \frac{Nq^2}{dVm\epsilon_\circ} \sum_{i=1}^{n} \frac{f_i}{(\omega_i^2 - \omega^2) + j(\gamma_i + \frac{2}{T_{\circ i}})\omega} \tag{1.13}$$

f_i es un factor que mide la importancia relativa de las distintas transiciones electrónicas. El aspecto de la variación de la parte real e imaginaria de la constante dieléctrica relativa con respecto a la pulsación se muestra en la fig. {1.10}.

Resulta evidente que el espectro de absorción $\alpha(\omega)$ y la característica dispersiva del medio $n(\omega)$ son funciones íntimamente relacionadas entre sí. Ambas derivan de una misma función compleja analítica $\chi(\omega)$, lo que garantiza que existe una relación causal entre ambas. En estas condiciones, se puede demostrar que hay dos relaciónes integrales [62]–[87] (relaciones de Kramers-Kronig) entre las funciones $\alpha(\omega)$ y $n(\omega)$. Como normalmente es más fácil medir la característica de atenuación, damos la expresión que relaciona el índice de refracción con respecto a la constante de atenuación:

$$n(\omega) = 1 + \frac{c_\circ}{\omega}\mathcal{P}\int_0^\infty \frac{\alpha(\omega')d\omega'}{\omega'^2 - \omega^2} \tag{1.14}$$

Donde la $\mathcal{P}$ indica la parte principal de la integral de Cauchy.

1.6. Propagación de pulsos luminosos

Tras estudiar los procesos de interacción luz–materia, estamos en condiciones de describir la propagación de un pulso luminosos por un medio material dieléctrico [3]. Este pulso tienen una única descomposición en ondas planas progresivas y monocromáticas (calculando la transformada de Fourier). Puesto que el medio es lineal (suponemos que las potencias ópticas son lo suficientemente bajas como para que los efectos no lineales sean despreciables), podremos propagar cada componente independientemente (empleando para ello un vector de onda ligeramente distinto para cada componente, ya que los coeficientes de atenuación e índice de refracción dependen de la pulsación) y volver a sumarlas tras la propagación (esto es, calcular la transformada inversa de Fourier). A partir de ahora vamos a denotar la pulsación de la portadora luminosa mediante $\omega_\circ$ porque es una notación que está muy extendida. Por supuesto, *no debemos confundirla con la pulsación propia de los dipolos*, ya que a la pulsación de trabajo el coeficiente de atenución de la fibra es muy bajo lo que implica que estamos muy lejos de ésta.

En la fig. {1.11} se muestra la dependencia temporal de los pulsos que se van a propagar: una portadora luminosa de una pulsación perfectamente determinada, $\omega_\circ$, modulada por una envolvente de duración finita. Si la respuesta impulsional del medio de propagación es $h(L,t)$, donde L es la distancia recorrida, resulta que podemos expresar el pulso propagado de dos maneras distintas: 1.- Como el producto de convolución del pulso por la respuesta impulsional del canal (o medio de propagación). 2.- Como la transformada inversa de Fourier del producto de la función de transferencia del canal multiplacada por la transformada de Fourier de la señal que se propaga:

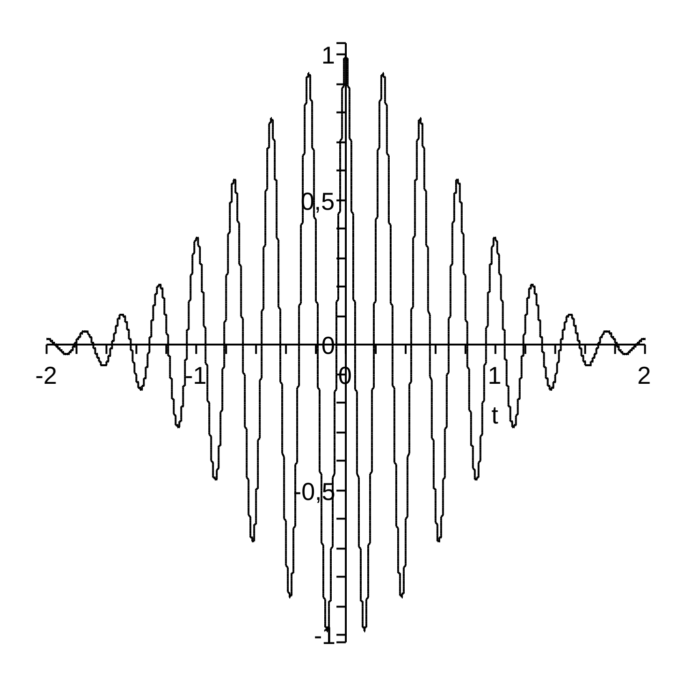


Figura 1.11: Pulso luminoso portador de información. En el eje vertical se ha representado el campo eléctrico normalizado, $\frac{E(x)}{E_\circ}$ y en el horizontal el tiempo normalizado $\frac{2\pi t}{T}t$.

$$\left.\begin{array}{r} U(0,t) = A(0,t)e^{j\omega_\circ t} \\ h(L,t) \end{array}\right\} \Rightarrow U(L,t) = U(0,t) * h(L,t) = \mathcal{F}^{-1}\{\mathcal{F}[A(0,t)e^{j\omega_\circ t}]\mathcal{F}[h(L,t)]\} =$$
$$\mathcal{F}^{-1}\{\mathcal{F}[A(0,t)e^{j\omega_\circ t}]e^{\frac{-\alpha(\omega)L}{2}}e^{-j\beta(\omega)L}\} = \mathcal{F}^{-1}\{\tilde{A}(0,\omega-\omega_\circ)e^{\frac{-\alpha(\omega)L}{2}}e^{-j\beta(\omega)L}\} \equiv$$
$$\equiv \mathcal{F}^{-1}\{\tilde{A}(0,\Delta\omega)e^{\frac{-\alpha(\omega)L}{2}}e^{-j\beta(\omega)L}\} \qquad (1.15)$$

En el desarrollo anterior hemos deducido que la transformada de Fourier del pulso a la entrada del canal es igual a la transformada de la envolvente, desplazada alrededor de la portadora luminosa $\omega_\circ$. Ahora vamos a tener en cuenta el orden de magnitud de la pulsación de la portadora de la información ($\omega_\circ \simeq 2\pi \times 192T\frac{rad}{s}$) y la duración típica de las señales luminosas (del orden de unos 50ps para las tasas binarias máximas). Esto hace que el ancho de banda asociado a la envolvente del pulso sea muy pequeña en comparación con $\omega_\circ$, por lo que podemos aproximar las funciones $e^{\frac{-\alpha(\omega)L}{2}}$ y $e^{-j\beta(\omega)L}$ por medio de un desarrollo en serie de Taylor alrededor de $\omega_\circ$. La dependencia de α con respecto a ω es muy débil y podemos considerar que es una constante. Tenemos, por tanto, que la función de transferencia del canal (medio de propagación) es aproximadamente igual a:

$$\left.\begin{array}{c} \alpha(\omega) \simeq \alpha_\circ \\ \beta(\omega) \simeq \beta_\circ + \dot{\beta}_\circ L(\omega-\omega_\circ) + \frac{\ddot{\beta}_\circ}{2}L(\omega-\omega_\circ)^2 \end{array}\right\} \Rightarrow$$
$$\left.\begin{array}{c} e^{\frac{-\alpha(\omega)L}{2}} \simeq e^{\frac{-\alpha_\circ L}{2}} \\ e^{-j\beta(\omega)L} \simeq e^{-j(\beta_\circ L + \dot{\beta}_\circ L\Delta\omega + \frac{\ddot{\beta}_\circ}{2}L\Delta\omega^2)} \end{array}\right\} \Rightarrow \left.\begin{array}{c} e^{\frac{-\alpha(\omega)L}{2}} \simeq e^{\frac{-\alpha_\circ L}{2}} \\ e^{-j\beta(\omega)L} \simeq e^{-j(\beta_\circ L + \dot{\beta}_\circ L\xi + \frac{\ddot{\beta}_\circ}{2}L\xi^2)} \end{array}\right\} \Rightarrow$$

$$\mathcal{H}(\xi, L) = e^{\frac{-\alpha_\circ L}{2}} e^{-j\beta_\circ L} e^{-j\dot{\beta}_\circ L\xi} e^{-j\frac{\ddot{\beta}_\circ}{2}L\xi^2} \qquad (1.16)$$
$$\xi := \omega - \omega_\circ \in (-\infty, +\infty)$$

Por tanto, el pulso propagado a lo largo del canal es:

$$U(L, t) = \mathcal{F}^{-1}\{\tilde{A}(0, \xi) e^{-j\dot{\beta}_\circ L\xi} e^{-j\frac{\ddot{\beta}_\circ}{2}L\xi^2}\} e^{\frac{-\alpha_\circ L}{2}} e^{j(\omega_\circ t - \beta_\circ L)} \qquad (1.17)$$

Vemos, pues, que al hacer el cambio de variable $\xi := \omega - \omega_\circ$, podemos trabajar con la transformada de Fourier de la envolvente compleja (que es una señal en banda base) y con una función de transferencia transladada del canal al origen de frecuencias (a banda base) que, en definitiva, veremos que no es más que un filtro de fase. Analicemos, a continuación, el significado físico de los distintos términos que aparecen en la expresión final:

- $e^{\frac{-\alpha_\circ L}{2}}$ es un término de atenuación, que afecta por igual a las distintas componentes espectrales. Todas ellas pierden energía en igual grado a medida que se van propagando por el canal por lo que el pulso va perdiendo amplitud (energía) sin sufrir ninguna otra distorsión en su forma temporal. Si hacemos caso omiso de este término, el canal se comporta como un filtro de fase.

- $e^{j(\omega_\circ t - \beta_\circ L)}$ nos permite recuperar la portadora de la señal, así como su referencia de fase si multiplicamos la envolvente compleja (particularizada en $z = L$) por este término (con lo que obtenemos la señal física).

- $e^{-j\dot{\beta}_\circ L\xi}$ es un término lineal en ξ del filtro de fase. En el dominio temporal representa un retardo puro de valor $t_r = \dot{\beta}_\circ L$. Es es tiempo que tarda el pulso luminoso en llegar al otro extremo del canal.

- $e^{-j\frac{\ddot{\beta}_\circ}{2}L\xi^2}$ es el término dispersivo y es el que en definitiva realmente nos interesa. Las distintas componentes espectrales tienen retardos de grupo ligeramente distintos a $t_r = \dot{\beta}_\circ L$ y estos retardos diferenciales están representados por este término. Por esta razón la expresión $\tau = \|\frac{\ddot{\beta}_\circ}{2}\|L\Delta\omega$ es una buena aproximación del ensanchamiento temporal que sufre el pulso tras recorrer la distancia L y donde $\Delta\omega$ es la anchura espectral del pulso que se propaga.

Si sólo estamos interesados en calcular la dispersión del pulso, podemos ignorar el retardo temporal y la atenuación, y tomar las siguientes expresiones como 1.- la función de transferencia del canal (se muestra en la fig. {1.12}) y 2.- la expresión del pulso dispersado:

$$\mathcal{H}(\xi, L) = e^{-j\frac{\ddot{\beta}_\circ}{2}L\xi^2}$$
$$U(L, t') = \mathcal{F}^{-1}\{\tilde{A}(0, \xi) e^{-j\frac{\ddot{\beta}_\circ}{2}L\xi^2}\} \qquad (1.18)$$

Por supuesto ésta no es una señal física, ya que se le ha eliminado la portadora. Antes de finalizar esta sección es importante que recordemos lo siguiente:

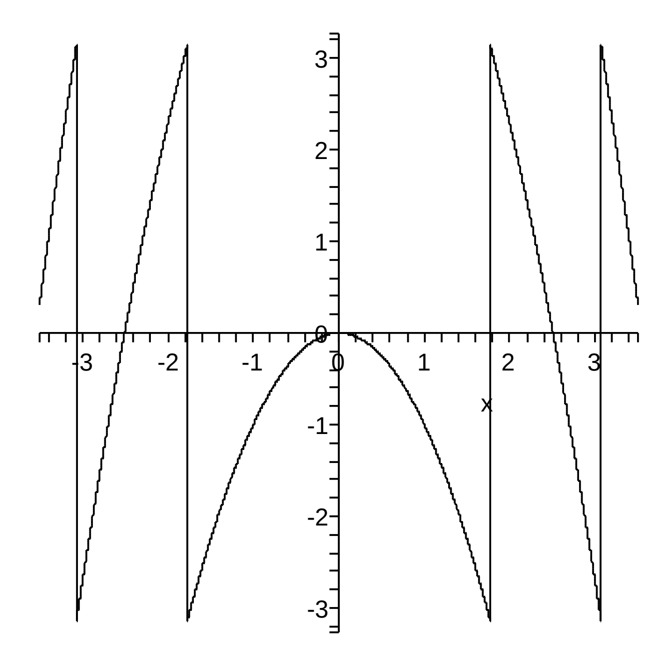


Figura 1.12: Argumento del filtro de fase cuadrático. El eje vertical representa el argumento en el rango $(-\pi, \pi)$ y el horizontal la pulsación adimensional $\frac{\xi}{\sqrt{\frac{2}{\beta_o}L}}$

.

1. $\tilde{A}(0, \xi)$ no es la transformada de Fourier del pulso luminoso a la entrada del canal. Es la transformada de Fourier de la *envolvente compleja* del pulso, que resulta de eliminar la portadora luminosa. En general ésta es una función compleja, ya que es común que los láseres con modulación directa emitan luz que sufre una ligera variación de la pulsación a lo largo de la duración la envolvente. Es el denominado *chirp* que se expresa matemáticamente como un argumento no constante de la envolvente compleja.

2. $A(L, t)$ no es el pulso a la salida del canal. Es su envolvente compleja. Para recuperar el pulso, tenemos que multiplicar por $e^{j(\beta_o L - \omega_o t)}$.

1.7. Propagación de un pulso de envolvente gaussiana

Como ejemplo vamos a considerar un pulso cuya pulsación es constante y cuya envolvente es gaussiana [26] (en este caso $\tilde{A}(0, \xi)$ es una función real). Veremos que este problema tiene solución analítica y nos servirá para comprobar que, además del fenómeno de dispersión (ensanchamiento temporal de la envolvente del pulso), se produce también una redistribución no uniforme de las componentes espectrales durante la propagación (es decir, el pulso adquiere un chirp). Apliquemos el método general deducido en el apartado anterior:

$$E(0,t) = A(0,t)e^{-j\omega_o t} = A_o e^{-\frac{t^2}{\tau_o^2}} e^{-j\omega_o t} \Rightarrow \mathcal{F}[A(0,t)] = \tilde{A}(0,\xi) = A_o \tau_o \sqrt{\pi} e^{-\frac{\xi^2}{(2/\tau_o)^2}} \quad (1.19)$$

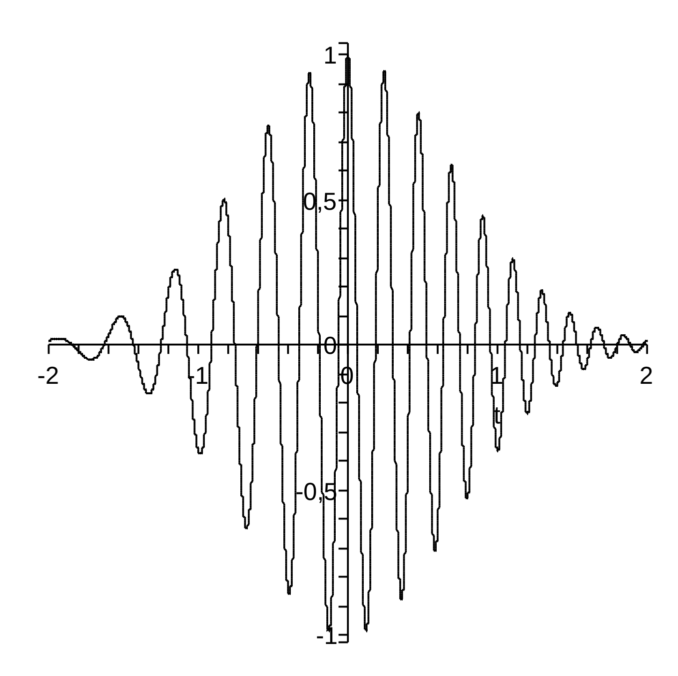


Figura 1.13: Pulso luminoso propagado por un medio con dispersión cromática positiva.En el eje vertical se ha representado el campo eléctrico normalizado y en el horizontal el tiempo normalizado $\frac{2\pi t}{T}$.

Puesto que la función de transferencia es:

$$\mathcal{H}(\xi, L) = e^{\frac{-\alpha_\circ L}{2}} e^{-j\dot{\beta}_\circ L\xi} e^{\frac{\ddot{\beta}_\circ}{2}L\xi^2} \tag{1.20}$$

La transformada de Fourier del pulso propagado es:

$$\begin{aligned} \tilde{A}(L,\xi) &= A_\circ \tau_\circ \sqrt{\pi} e^{\frac{-\alpha_\circ L}{2}} e^{-j\dot{\beta}_\circ L\xi} e^{-\frac{\xi^2}{4}(\tau_\circ^2 + 2j\ddot{\beta}_\circ L)} = \\ &= \frac{A_\circ \tau_\circ \sqrt{\pi}}{\sqrt{\tau_\circ^2 + 2j\ddot{\beta}_\circ L}} e^{\frac{-\alpha_\circ L}{2}} e^{-j\dot{\beta}_\circ L\xi} \sqrt{\tau_\circ^2 + 2j\ddot{\beta}_\circ L}\, exp\left(-\frac{\xi^2}{\frac{2^2}{\tau_\circ^2 + 2j\ddot{\beta}_\circ L}}\right) \end{aligned} \tag{1.21}$$

En la última igualdad hemos multiplicado y dividido por la desviación típica de la función Gaussian compleja para poder calcular la transformada inversa de Fourier. El término en $exp[-j\dot{\beta}_\circ L\xi]$ representa un retardo temporal puro de $\dot{\beta}_\circ L$ segundos y no nos interesa especialmente (ya sabemos que la velocidad de grupo del pulso es $\frac{1}{\dot{\beta}_\circ}$), por lo que definimos un nuevo tiempo $t' = t - \dot{\beta}_\circ L$ y expresamos el pulso propagado en función de esta nueva variable temporal:

$$A(L, t') = \mathcal{F}^{-1}[\tilde{A}(L,\xi)] = \frac{A_\circ \tau_\circ}{\sqrt{\tau_\circ^2 + 2j\ddot{\beta}_\circ L}} e^{-\frac{\alpha_\circ L}{2}} e^{-\frac{t'^2}{\tau_\circ^2 + 2j\ddot{\beta}_\circ L}} \tag{1.22}$$

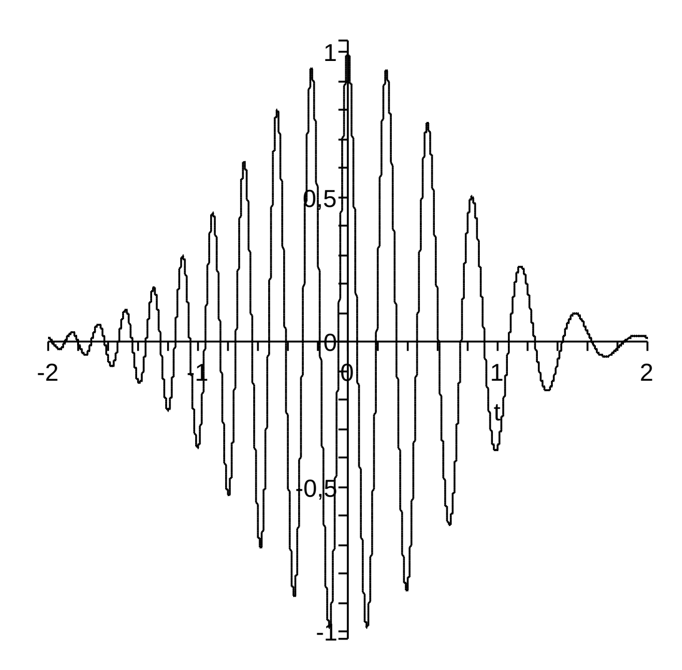


Figura 1.14: Pulso luminoso propagado por un medio con dispersión cromática negativa. En el eje vertical se ha representado el campo eléctrico normalizado $\frac{E(t)}{E_\circ}$ y en el horizontal el tiempo normalizado $\frac{2\pi t}{T}$.

Para poder interpretar qué ha ocurrido con el pulso después de propagarse, tenemos que reordenar la expresión anterior, de manera que se vea de una manera clara cuál es la envolvente y cuál es la fase. Para ello descomponemos el argumento de la función exponencial como la suma de su parte real más su parte imaginaria y el coeficiente primero lo expresamos en forma polar forma polar (módulo–fase):

$$\frac{t'^2}{\tau_\circ^2 + 2j\ddot{\beta}_\circ L} = \frac{t'^2}{\tau_\circ^2 + 2j\ddot{\beta}_\circ L}\frac{\tau_\circ^2 - 2j\ddot{\beta}_\circ L}{\tau_\circ^2 - 2j\ddot{\beta}_\circ L} = \frac{t'^2}{\tau_\circ^2 + 4\dfrac{\ddot{\beta}_\circ^2 L^2}{\tau_\circ^2}} - j\frac{2\ddot{\beta}_\circ L}{\tau_\circ^4 + 4\ddot{\beta}_\circ^2 L^2}t'^2$$

$$\sqrt{\tau_\circ^2 + 2j\ddot{\beta}_\circ L} = \sqrt{\sqrt{\tau_\circ^4 + 4\ddot{\beta}_\circ^2 L^2}exp\left[j\ \mathrm{tg}^{-1}\left(\frac{2\ddot{\beta}_\circ L}{\tau_\circ^2}\right)\right]} =$$

$$= \sqrt[4]{\tau_\circ^4 + 4\ddot{\beta}_\circ^2 L^2}exp\left[\frac{j}{2}\ \mathrm{tg}^{-1}\left(\frac{2\ddot{\beta}_\circ L}{\tau_\circ^2}\right)\right] \tag{1.23}$$

Teniendo en cuenta las relaciones anteriores, resulta:

$$A(L,t') = \frac{A_\circ \tau_\circ}{\sqrt[4]{\tau_\circ^4 + 4\ddot{\beta}_\circ^2 L^2}}e^{-\dfrac{\alpha_\circ L}{2}} exp\left(-\frac{t'^2}{\tau_\circ^2 + 4\frac{\ddot{\beta}_\circ^2 L^2}{\tau_\circ^2}}\right) exp\ j\left[-\frac{1}{2}\ \mathrm{tg}^{-1}\left(\frac{2\ddot{\beta}_\circ L}{\tau_\circ^2}\right) + \frac{2\ddot{\beta}_\circ L t'^2}{\tau_\circ^4 + 4\ddot{\beta}_\circ^2 L^2}\right] \tag{1.24}$$

Interpretar esta fórmula es sencillo, ya que podemos determinar la expresión de la envolvente y de la frecuencia instantánea:

$$\begin{aligned} \|A(L,t')\| &= exp\left(\frac{-\alpha_\circ L}{2}\right)\frac{A_\circ\tau_\circ}{\sqrt[4]{\tau_\circ^4+4\ddot{\beta}_\circ^2L^2}}exp\left(-\frac{t'^2}{\tau_\circ^2+4\frac{\ddot{\beta}_\circ^2L^2}{\tau_\circ^2}}\right) \\ \arg[A(L,t')] &= -\frac{1}{2}\,\mathrm{tg}^{-1}\left(\frac{2\ddot{\beta}_\circ L}{\tau_\circ^2}\right)+\frac{2\ddot{\beta}_\circ L}{\tau_\circ^4+4\ddot{\beta}_\circ^2L^2}t'^2 \Rightarrow \\ \xi(t') &:= \frac{1}{2\pi}\frac{\partial\arg[A(L,t')]}{\partial t'}=\frac{2}{\pi}\frac{\ddot{\beta}_\circ L}{\tau_\circ^4+4\ddot{\beta}_\circ^2L^2}t' \end{aligned} \tag{1.25}$$

En estas expresiones comprobamos que la desviación típica de la envolvente del pulso crece linealmente con la distancia recorrida y con el valor absoluto del coeficiente de dispersión cromática. Advertimos, además, que en el pulso se produce un cambio en la pulsación instantánea, aumentando ésta linealmente con el tiempo (adquiere un chirp positivo) si el signo del coeficiente de dispersión cromática es positivo y disminuyendo (adquiere un chirp negativo) si el signo del coeficiente de dispersión cromática es negativo, tal como se muestra en las fig. {1.13} y {1.14}. Si queremos pasar a frecuencias físicas, no tenemos más que sustituir $\omega = \omega_\circ + \xi$.

1.8. Problemas

Problema 1: Determinar el ensanchamiento temporal que sufre un pulso de envolvente gaussiana si está inicialmente chirpeado. Considere los cuatro casos posibles: chirp inicial positivo o negativo y dispersión cromática positiva o negativa.

Solución: La envolvente compleja de un pulso con chirp tiene la siguiente expresión [26] analítica[10]:

$$A(0,t) = A_\circ e^{-\frac{(1+jC_\circ)t^2}{2\tau_\circ^2}} \Rightarrow \mathcal{F}[A(0,t)] = \tilde{A}(0,\xi) = A_\circ\sqrt{\frac{2\pi\tau_\circ^2}{1+jC_\circ}} exp\left(-\frac{\xi^2}{\frac{2(1+jC_\circ)}{\tau_\circ^2}}\right)$$

Vemos que la desviación típica del pulso se ve incrementada por un factor $\sqrt{1+C_\circ^2}$ frente al caso no chirpeado. $C_\circ$ representa la pendiente de la pulsación a lo largo del pulso cambiada de signo, ya que si derivamos la parte compleja del argumento de la función gaussiana obtenemos: $\xi = -\frac{C_\circ t}{\tau_\circ^2}$. La transformada de Fourier del pulso propagado es:

$$\tilde{A}(L,\xi) = A_\circ\sqrt{\frac{2\pi\tau_\circ^2}{1+jC_\circ}} exp\left[-\frac{\xi^2}{\frac{2(1+jC_\circ)}{\tau_\circ^2}}\right] exp\left(-j\frac{\ddot{\beta}_\circ}{2}L\xi^2\right) =$$

$$= A_\circ\sqrt{\frac{2\pi\tau_\circ^2}{1+jC_\circ}} exp\left[-\xi^2\frac{\tau_\circ^2+(1+jC_\circ)j\ddot{\beta}_\circ L}{2(1+jC_\circ)}\right] \Rightarrow$$

$$\tilde{A}(L,\xi) = A_\circ\tau_\circ\sqrt{\frac{1}{\tau_\circ^2+(1+jC_\circ)j\ddot{\beta}_\circ L}} exp\left[-\frac{\xi^2}{\frac{4}{\frac{2\tau_\circ^2+2(1+jC_\circ)j\ddot{\beta}_\circ L}{1+jC_\circ}}}\right]\sqrt{\frac{\pi(2\tau_\circ^2+2(1+jC_\circ)j\ddot{\beta}_\circ L)}{1+jC_\circ}}$$

Si ahora calculamos la transformada inversa obtenemos:

$$A(L,t^{'}) = \mathcal{F}^{-1}[\tilde{A}(L,\xi)] = \frac{A_\circ\tau_\circ}{\sqrt{\tau_\circ^2+(1+jC_\circ)j\ddot{\beta}_\circ L}} exp\left(-\frac{t'^2(1+jC_\circ)}{2\tau_\circ^2+2(1+jC_\circ)j\ddot{\beta}_\circ L}\right)$$

Si ahora tenemos en cuenta que:

$$\frac{A_\circ\tau_\circ}{\sqrt{\tau_\circ^2+(1+jC_\circ)j\ddot{\beta}_\circ L}} = \frac{A_\circ\tau_\circ}{\sqrt[4]{(1-\frac{C_\circ\ddot{\beta}_\circ L}{\tau_\circ^2})^2+\frac{\ddot{\beta}_\circ^2 L^2}{\tau_\circ^4}}} exp\left[-\frac{j}{2}tg^{-1}\left(\frac{\ddot{\beta}_\circ L}{\tau_\circ^2-C_\circ\ddot{\beta}_\circ L}\right)\right]$$

y que:

[10] No estamos siendo consistentes con las expresiones utilizadas anteriormente. Aquí $\tau_\circ$ es la desviación típica.

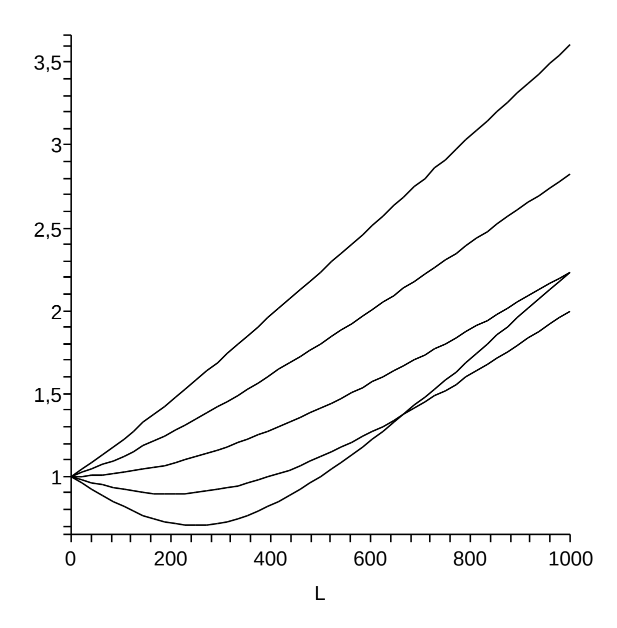


Figura 1.15: En esta figura se ha representado el ensanchamiento sufrido por los pulsos, $\frac{\tau}{\tau_\circ}$, frente a la distancia recorrida en Km. El parámetro dispersivo es $\ddot{\beta} = 20ps^2/Km$, la duración inicial de los pulsos era de 100ps y el parámetro del chirp $C_\circ$ ha tomado los valores -1, -0.5, 0, 0.5 y 1.

$$-\frac{t'^2}{2}\frac{(1+jC_\circ)}{(\tau_\circ^2 - C_\circ\ddot{\beta}_\circ L) + j\ddot{\beta}_\circ L} = -\frac{t'^2}{2}\frac{(1+jC_\circ)}{(\tau_\circ^2 - C_\circ\ddot{\beta}_\circ L) + j\ddot{\beta}_\circ L}\frac{(\tau_\circ^2 - C_\circ\ddot{\beta}_\circ L) - j\ddot{\beta}_\circ L}{(\tau_\circ^2 - C_\circ\ddot{\beta}_\circ L) - j\ddot{\beta}_\circ L} =$$
$$-\frac{t'^2}{2}\frac{\tau_\circ^2 + j[C_\circ\tau_\circ^2 - (1+C_\circ^2)\ddot{\beta}_\circ L]}{(\tau_\circ^2 - C_\circ\ddot{\beta}_\circ L)^2 + \ddot{\beta}_\circ^2 L^2}$$

Obtenemos finalmente:

$$A(L,t') = \frac{A_\circ}{\sqrt[4]{(1-\frac{C_\circ\ddot{\beta}_\circ L}{\tau_\circ^2})^2 + \frac{\ddot{\beta}_\circ^2 L^2}{\tau_\circ^2}}} exp\left[-\frac{1}{2}\frac{t'^2}{\tau_\circ^2 + \frac{\ddot{\beta}_\circ^2 L^2}{\tau_\circ^2}(1+C_\circ^2) - 2C_\circ\ddot{\beta}_\circ L}\right]$$
$$exp\left[-\frac{t'^2}{2}\frac{j[C_\circ\tau_\circ^2 - (1+C_\circ^2)\ddot{\beta}_\circ L]}{(\tau_\circ^2 - C_\circ\ddot{\beta}_\circ L)^2 + \ddot{\beta}_\circ^2 L^2} - \frac{j}{2}tg^{-1}\left(\frac{\ddot{\beta}_\circ L}{\tau_\circ^2 - C_\circ\ddot{\beta}_\circ L}\right)\right]$$

Ya estamos en disposición de identificar el módulo y el argumento de la función:

$$\|A(L,t')\| = \frac{A_\circ}{\sqrt[4]{(1-\frac{C_\circ\ddot{\beta}_\circ L}{\tau_\circ^2})^2 + \frac{\ddot{\beta}_\circ^2 L^2}{\tau_\circ^2}}} exp\left[-\frac{1}{2}\frac{t'^2}{\tau_\circ^2 + \frac{\ddot{\beta}_\circ^2 L^2}{\tau_\circ^2}(1+C_\circ^2) - 2C_\circ\ddot{\beta}_\circ L}\right]$$

$$\arg[A(L,t')] = -\frac{t'^2}{2}\frac{j[C_\circ\tau_\circ^2 - (1+C_\circ^2)\ddot{\beta}_\circ L]}{(\tau_\circ^2 - C_\circ\ddot{\beta}_\circ L)^2 + \ddot{\beta}_\circ^2 L^2} - \frac{j}{2}tg^{-1}\frac{\ddot{\beta}_\circ L}{\tau_\circ^2 - C_\circ\ddot{\beta}_\circ L}$$

Viendo de la ecuación (1.26) es fácil comprender el comportamiento cualitativo de la propagación de un pulso chirpeado. Siempre que la dispersión cromática ($\ddot{\beta}$) y el coeficiente asociado al chirp ($C_\circ$) tengan signos iguales, existirá una determinada distancia durante la que el pulso se comprime y, a partir de ese momento, se dispersa a un ritmo que ya no depende del signo del chirp, sino de su valor absoluto. Esto se muestra en la fig. {1.15}. Se puede, además, calcular la distancia a la cual la duración del pulso es mínima; basta con minimizar, en función de L, la desviación cuadrática media del pulso:

$$\frac{\partial}{\partial L}\left[\tau_\circ^2 + \frac{\ddot{\beta}_\circ^2 L^2}{\tau_\circ^2}(1 + C_\circ^2) - 2C_\circ\ddot{\beta}_\circ L\right] = 0 \Rightarrow L = \frac{C_\circ}{1 + C_\circ^2}\frac{\tau_\circ^2}{\ddot{\beta}_\circ}$$

De nuevo podemos comprobar que sólo se obtiene una solución con sentido físico si el signo del chirp y de la dispersión cromática son iguales. Además vemos que para esta distancia se anula el chirp en el pulso, con lo que las condiciones de compresión máxima y chirp nulo coinciden:

$$\frac{\partial}{\partial t'}\arg[A(L, t')] = 0 \Rightarrow L = \frac{C_\circ}{1 + C_\circ^2}\frac{\tau_\circ^2}{\ddot{\beta}_\circ}$$

Problema 2: Considere la propagación de un pulso de envolvente gaussiana. La frecuencia es de 192THz, la duración inicial es $\tau_\circ = 1ps$ y el coeficiente de dispersión cromática es $\ddot{\beta} = 20ps^2/Km$. El pulso recorre una distancia de 100Km. Se pide estimar las frecuencias máxima y mínima del pulso y su posición relativa.

Solución: Tal como hemos visto, la desviación típica de un pulso gaussiano no chirpeado crece a un ritmo fijado por la siguiente expresión:

$$\tau_o^{'} = \sqrt{\tau_o^2 + 4\frac{\ddot{\beta}_o^2 L^2}{\tau_o^2}} \simeq 2\frac{\ddot{\beta}_o L}{\tau_o}$$

donde hemos tenido en cuenta que el ensanchamiento temporal producido a lo largo de los 100Km es grande en comparación con la duración inicial del pulso. Vamos a considerar como extremos temporales del pulso ensanchado, los tiempos $t = -3\tau_\circ'$ y $t = 3\tau_\circ'$.

Las excursiones en frecuencia (frente a 192THz) están dadas por:

$$\xi(t^{'}) = \frac{2}{\pi}\frac{\ddot{\beta}_o L}{\tau_o^4 + 4\ddot{\beta}_o^2 L^2} t^{'} \simeq \frac{1}{2\pi\ddot{\beta}_o L} t^{'}$$

con lo cual:

$$\begin{aligned}
\Delta\tau_o^{'} &= 6\tau_o^{'} = 12\frac{20ps^2/Km \times 100Km}{1ps} = 24ns \\
f_{max} &= f_\circ + \frac{3}{\pi\tau_\circ} = 192,955THz \\
f_{min} &= f_\circ - \frac{3}{\pi\tau_\circ} = 191,045THz
\end{aligned}$$

Como el signo de la dispersión cromática es positivo, eso indica que el timepo de propagación crece con la frecuencia, por lo que primero llegan las freceuncias más bajas y luego las más altas.

Capítulo 2

Propagación de ondas luminosas en fibras

Resumen: Para describir el comportamiento de una fibra óptica, necesitamos conocer tres propiedades de ésta: 1.- La característica de inyección, 2.- La característica de atenuación y 3.- La característica dispersiva. Las características de atenuación y dispersión dependen, fundamentalmente, de las propiedades físicas del material con el que se fabrican las fibras (generalmente sílice), aunque en la dispersión aparece un término adicional asociado al guiado de la luz. En fibras que soportan más de un modo de propagación (fibras multimodo) este término recibe el nombre de dispersión intermodal, mientras que en las que existe un único modo (fibras monomodo) se le llama dispersión de guía.
En este capítulo estudiaremos superficialmente las fibras multimodo (usando la aproximación de la óptica geométrica), mientras que analizaremos más detalladamente la propagación de luz en fibras monomodo utilizando la óptica electromagnética. El desarrollo matemático que permite deducir su ecuación de dispersión sólo se va a esbozar, dejándose como ejercicio para el lector que complete los pasos omitidos. En el próximo capítulo emplearemos los resultados obtenidos en éste para determinar las características dispersivas de una fibra monomodo estándar.

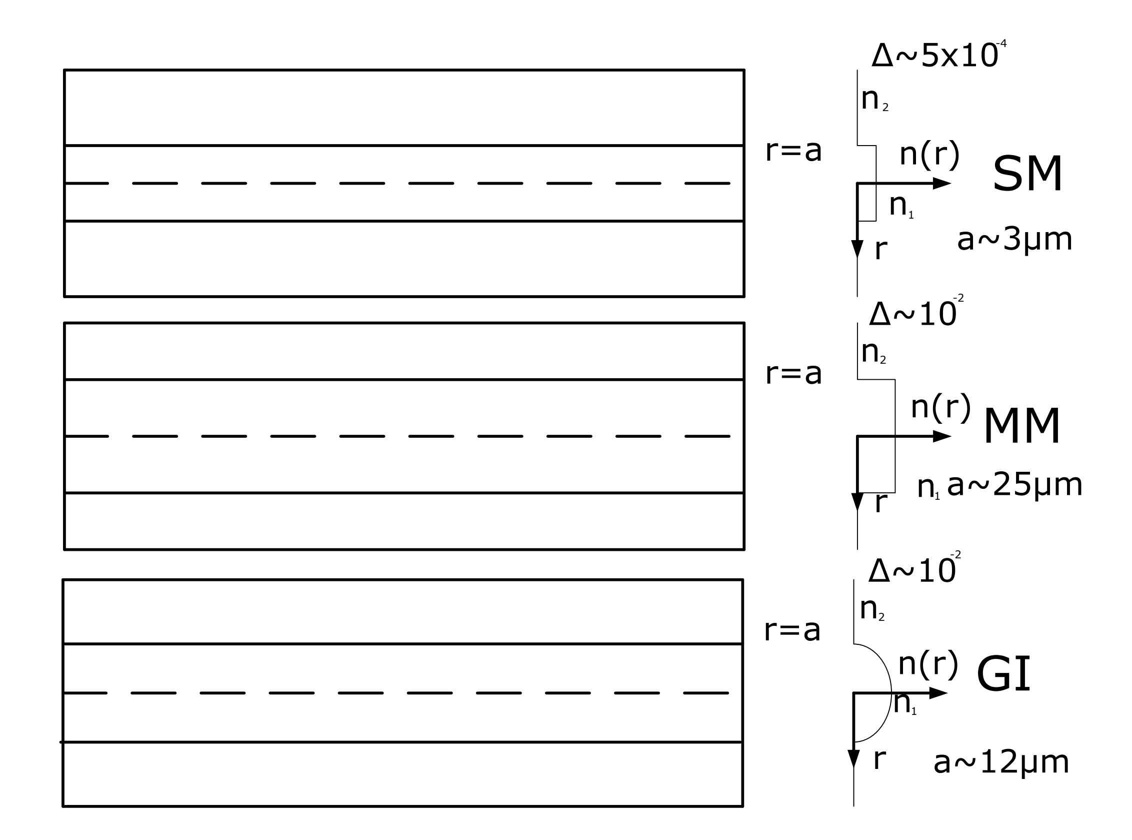


Figura 2.1: Tres tipos de fibras ópticas; a representa el radio del núcleo y Δ el salto de índice adimensional.

2.1. Introducción

La fibra óptica permite guiar longitudinalmente los pulsos luminosos, evitando que su energía se diperse. El guiado de la luz se consigue gracias a que en las cercanías del eje de la fibra hay una región cilíndrica, donde el índice de refracción es más elevado: esta zona de la fibra se denomina núcleo. El resto recibe el nombre de corteza e, idealmente, se extiende transversalmente hasta el infinito. Las fibras ópticas se pueden clasificar en dos grandes familias, según soporten muchos modos [1] de propagación (fibras multimodo) [2] o un único modo de propagación (fibras monomodo).

En la fig. {2.1} se muestra la estructura optogeómetrica de tres tipos de fibras: la fibra multimodo a salto de índice (MM), la fibra multimodo a gradiente de índice (GI) y la fibra monomodo estándar (SM). Vemos que las fibras multimodo tienen un núcleo de diámetro grande y una diferencia relativamente importante entre los índices de núcleo y corteza. Por contra, la fibra monodo estándar tiene un núcleo de diámetro pequeño y la diferencia entre los índices de núcleo y corteza es mucho menor que en las multimodo. De esta manera se consigue, tal y como probaremos más tarde, que, a la longitud de onda de trabajo, exista un único modo guiado.

[1] Un modo de propagación es una distribución transversalmente estacionaria y longitudinalmente progresiva del campo electromagnético. Además, la densidad de energía debe tender a cero a medida que nos alejamos del eje de la fibra.

[2] Típicamente varios centenares.

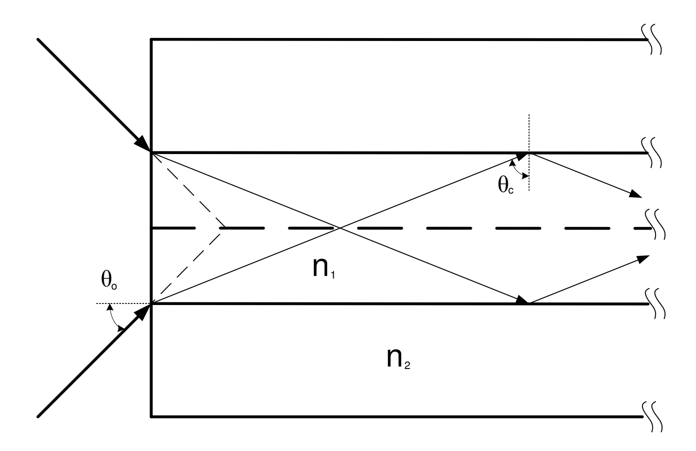


Figura 2.2: Apertura numérica de una fibra óptica.

De una fibra óptica nos interesa conocer tres propiedades distintas, que son:

1. **La característica de inyección**. Nos indica la forma en la que debemos proceder para inyectar de manera eficiente potencia óptica dentro de la fibra y conseguir, además, que ésta se propague longitudinalmente.

2. **La característica de atenuación**. La contante de atenuación permite determinar a qué ritmo pierden energía los pulsos luminosos durante su propagación por la fibra.

3. **La característica de dispersión**. Predice el ensanchamiento temporal que sufren los pulsos luminosos en función de la distancia recorrida a lo largo de la fibra.

2.2. Óptica geométrica

El uso de la óptica geométrica está justificado en aquellas situaciones en los que las dimensiones transversales del medio de propagación son muy grandes en comparación con la longitud de onda de la luz [17]. Las fibras multimodo a salto de índice tienen diámetros de núcleo del orden de $50\mu m$ y las multimodo a variación gradual del orden de $25\mu m$. En cambio las fibras monomodo estándar tienen núcleos mucho más pequeños: el diámetro es del orden de $5\mu m$. Así, pues, el uso de esta aproximación está relativamente justificado en las fibras multimodo, mientras que no lo está en absoluto en las fibras monomodo.

En la fig. {2.2} se ha representado la superficie de entrada de una fibra a salto de índice (es decir, una fibra en la que la variación de índice entre núcleo y corteza ocurre de manera brusca). ¿Qué tiene que ocurrir para que el rayo inyectado a la entrada se propague sin perder energía? Evidentemente, tiene que producirse una reflexión total en la superficie de separación entre el núcleo y la corteza. Es decir, se tiene que verificar que $\theta_r \geq \frac{\pi}{2}$, donde θ_r representa el ángulo que forma el rayo refractado (no representado en la figura) con respecto

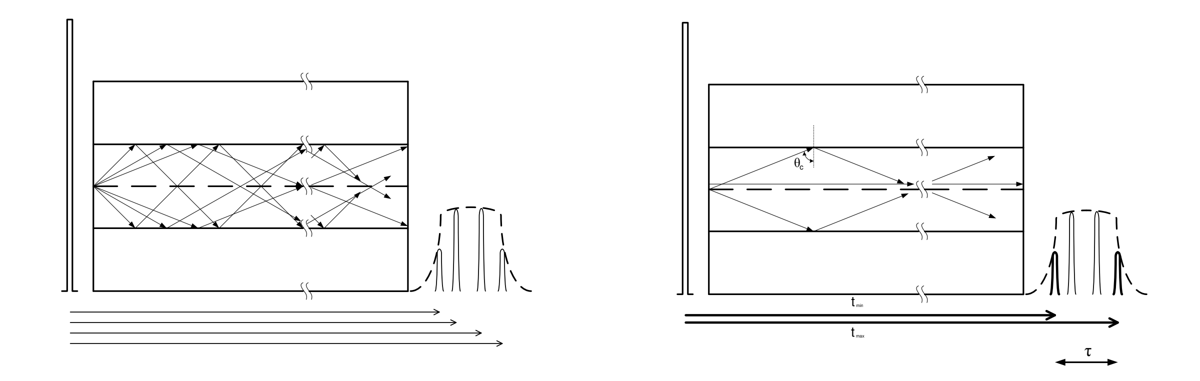


Figura 2.3: Dispersión intermodal en una fibra multimodo a salto de índice (izquierda) y cálculo de la disperisión asociada a los rayos de velocidades extremas (derecha).

a la normal a la superficie de separación entre núcleo y corteza. Para deducir qué rayo a la entrada cumple las condiciones críticas de incidencia (es decir, $\theta_r = \frac{\pi}{2}$), aplicamos la ley de Snell:

$$\left.\begin{array}{r} n_1 \sin\theta_i = n_2 \sin\theta_r \\ \theta_r = \frac{\pi}{2} \end{array}\right\} \Rightarrow \quad \sin\theta_i = \tfrac{n_2}{n_1} \Rightarrow n_\circ \sin\theta_\circ = n_1 \sin(\tfrac{\pi}{2} - \theta_i) = n_1 \cos\theta_i \quad \Rightarrow$$

$$\sin\theta_\circ = n_1 \cos\theta_i = n_1 \sqrt{1 - \tfrac{n_1^2}{n_2^2}} \tag{2.1}$$

Si ahora definimos:

$$\Delta := \frac{n_1^2 - n_2^2}{2n_1^2} = \frac{(n_1 + n_2)(n_1 - n_2)}{2n_1^2} \cong \frac{2n_1(n_1 - n_2)}{2n_1^2} = \frac{n_1 - n_2}{n_1} \tag{2.2}$$

En la expresión anterior hemos tenido en cuenta que los índices de refracción de núcleo y corteza toman valores muy próximos entre sí, de manera que a la hora de calcular $n_1 + n_2$ no cometemos un error apreciable si consideramos que ambos índices son iguales. La constante Δ recibe el nombre de *salto de índice adimensional* y la expresión que normalmente se utiliza para calcularlo es la última que es, además, la que justifica este nombre. Finalmente nos queda:

$$\sin\theta_\circ = n_1\sqrt{2\Delta} \Rightarrow \theta_\circ = \sin^{-1}(n_1\sqrt{2\Delta}) \simeq n_1\sqrt{2\Delta} \tag{2.3}$$

La última expresión es aproximadamente cierta debido a que el salto de índice adimensional es muy pequeño para todas las fibras utilizadas en comunicaciones ópticas. Por supuesto, el ángulo $\theta_\circ$ (que recibe el nombre de apertura numérica) se mide en unidades naturales, esto es, en radianes [3]. Ahora vamos a dar valores típicos para una fibra multimodo a

[3] En el razonamiento hecho, hemos supuesto que los rayos son meridianos, esto es, rayos que cortan el eje

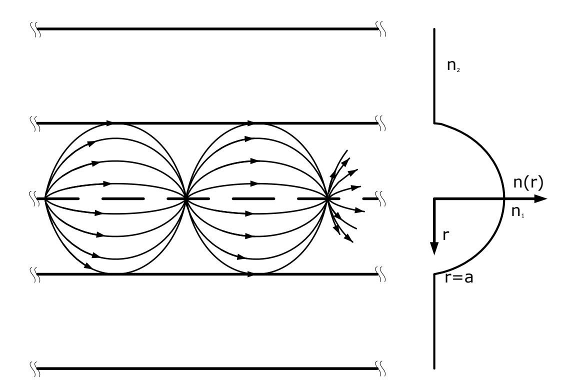


Figura 2.4: Igualación de los tiempos de propagación en una fibra multimodo a gradiente de índice.

salto de índice y para una fibra monomodo estándar. El resultado obtenido será cuantitativamente válido para fibras multimodo y, como mucho, sólo cualitativamente válido para fibras monomodo (lo que pretendemos es poner de manifiesto lo complicado que resulta acoplar luz en una fibra monomodo).

- **Fibra monomodo:** $n_1 \simeq 1,45, \quad \Delta \simeq 5 \times 10^{-4} \Rightarrow \theta_\circ \sim 3°$
- **Fibra multimodo:** $n_1 \simeq 1,45, \quad \Delta \simeq 1 \times 10^{-2} \Rightarrow \theta_\circ \sim 12°$

Desde el punto de vista de la característica de inyección, la fibra multimodo tiene mejores prestaciones que la monomodo, ya que resulta mucho más fácil acoplar potencia óptica en ellas. Además, dado que tienen núcleos muchos más grandes, resultan más fáciles de empalmar (ya sea mecánicamente o por soldadura), motivo por el cual los primeros sistemas de comunicaciones ópticas que se diseñaron empleaban este tipo de fibra.

La apertura numérica define el llamado *cono de aceptación de la fibra.* Cualquier rayo inyectado en el cono y que tenga un ángulo de incidencia inferior a la apertura numérica se propaga longitudinalmente. El cono está definido por la sección circular del núcleo y las rectas que, pasando por su circunferencia, cortan al eje formando un ángulo $\theta_\circ$, tal y como se miestra en la fig. {2.2}.

Ahora vamos a evaluar la característica dispersiva de las fibras multimodo a salto de índice, considerando sólo la propagación de rayos meridianos. En la fig. {2.3}, en su parte izquierda, se muestran distintas familias de rayos que, evidentemente, forman ángulos distintos con el eje de la fibra. Como la velocidad de la luz es, en todos los casos $\frac{c_\circ}{n_1}$, deducimos que la velocidad axial es distinta en cada familia de rayos. Por consiguiente, rayos que se inyectaron simultáneamente al principio de la fibra, llegan al otro extremo en instantes distintos, produciéndose un ensanchamiento temporal del pulso que se propaga por la fibra. Este ensanchamiento temporal se conoce con el nombre de *dispersión intermodal.* Vamos, ahora, a tratar de cuantificarlo.

---

longitudinal de la fibra. Si el rayo es oblicuo, se puede demostrar que su condición de reflexión total es más laxa, por lo que tenemos garantizado que si un rayo está dentro del cono de aceptación, se va a propagar longitudinalmente.

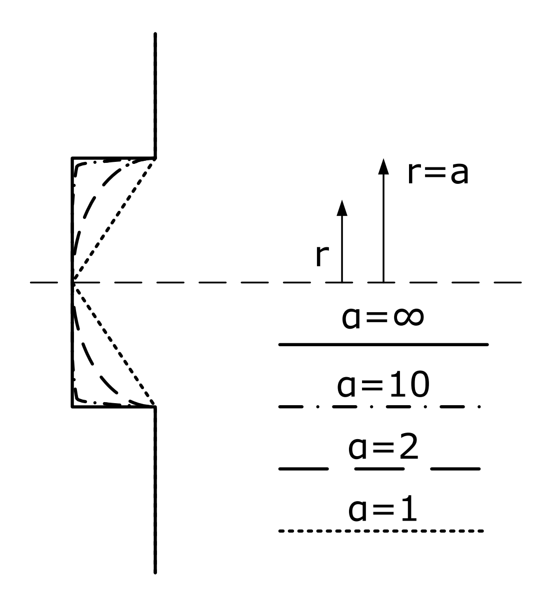


Figura 2.5: Perfil de variación del índice de refracción para varios valores de α.

En la misma fig. {2.3}, en su parte derecha, se ha representado también el rayo más veloz de todos (que se propaga paralelamente al eje, siguiendo la trayectoria más corta) y el más lento, para el que se dan condiciones críticas de reflexión y, por tanto, sigue la trayectoria más larga. Podemos calcular fácilmente el tiempo que tardan uno y otro en llegar al extremo de la fibra y, al calcular la diferencia, estimar el ensanchamiento temporal que sufre el pulso luminoso:

$$\left.\begin{array}{l} \frac{t_{max}}{L} = \frac{n_1}{c_\circ}\frac{1}{\sin(\theta_{cr})} = \frac{n_1^2}{c_\circ n_2} \\ \frac{t_{min}}{L} = \frac{n_1}{c_\circ} \end{array}\right\} \Rightarrow \frac{\tau}{L} = \frac{t_{max}-t_{min}}{L} = \frac{n_1}{c_\circ}(\frac{n_1}{n_2}-1) = \frac{n_1^2}{c_\circ n_2}\frac{n_1-n_2}{n_1} \Rightarrow$$

$$\frac{\tau}{L} = \frac{n_1^2\Delta}{c_\circ n_2} \tag{2.4}$$

Si damos valores típicos para el caso de una fibra multimodo a salto de índice (MM) tendremos que $n_1 \simeq 1,45$ y $\Delta \simeq 10^{-2}$ por lo que deducimos que $B_{max}L_{max} \simeq \frac{L}{\tau} \sim 5Mb/s \times Km$. Vemos, pues, que con este tipo de fibras no se pueden transmitir grandes tasas binarias a grandes distancias, motivo por el que no se usan comercialmente. La dispersión intermodal aparece como consecuencia de que existen distintas familias de rayos (modos de propagación) que viajan con velocidades longitudinales muy distintas. Por motivos obvios, en las fibras monomodo no hay dispersión intermodal.

Una forma de paliar este problema consiste en fabricar fibras multimodo en las que el índice de refracción no permanezca constante en el núcleo de la fibra. El índice disminuye a medida que nos alejamos del eje, de manera que aquellos rayos que tienen que recorrer distancias mayores (que se alejan más del eje de la fibra), tienen velocidades medias más altas ("ven", en promedio, índices menores) de forma que se reducen las diferencias entre

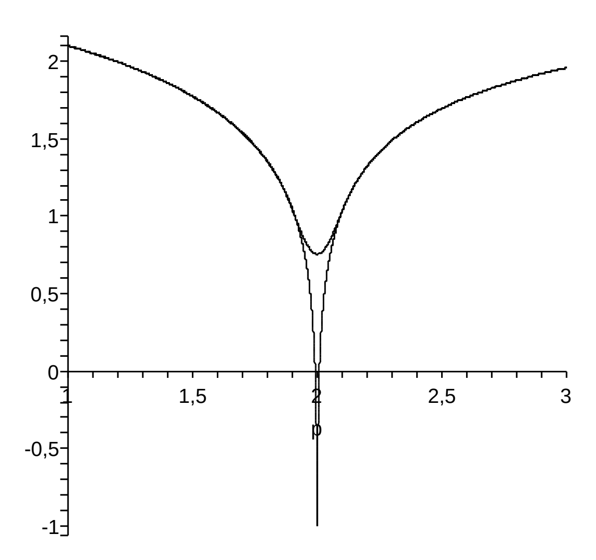


Figura 2.6: Dispersión intermodal en una fibra a gradiente de índice. Se muestran tanto la curva teórica como tecnológica. En el eje vertical se ha representado $\log_{10}(\frac{\tau/L}{1ns/Km})$ y en el horizontal el parámetro de definición del perfil que aquí se denota mediante la letra p.

los tiempos de propagación de los distintos modos guiado. Esto se muestra en la fig. {2.4}. Ahora se plantea un problema obvio. ¿Cuál es el perfil de variación del índice de refracción que minimiza la dispersión intermodal? Para resolver esta cuestión, definimos una familia de perfiles de variación del índice de refracción que depende de un único parámetro α, parámetro que al aumentar hace que la variación del índice entre el eje y la corteza sea más brusca. El rango de variación es $0 \leq \alpha \leq \infty$ y en la fig. {2.5} se muestran los perfiles asociados a algunos valores particulares de α. La expresión analítica de esta familia de perfiles es la siguiente:

$$n(r) := \begin{cases} n_1[1 - \Delta(\frac{r}{a})^{\alpha}] & r \leq a \\ n_1[1 - \Delta] = n_2 & r \geq a \end{cases} \tag{2.5}$$

Se puede calcular (aunque no lo haremos) la dispersión intermodal para cada valor del parámetro α y representar, por tanto, el valor de la dispersión frente a α, tal como se muestra en la fig. {2.6}.

La mejora que se obtiene [41]–[76] para el perfil óptimo (que es aproximadamente parabólico [4]) es extremadamente grande, siendo la expresión del ensanchamiento temporal del pulso la siguiente:

$$\frac{\tau}{L} = \frac{n_2\Delta^2}{8c_\circ} \tag{2.6}$$

[4]El perfil óptimo se produce para un valor de $\alpha = 2(1 - \Delta)$, aunque no lo vamos a demostrar. El lector interesado en conocer la justificación de este resultado puede consultar la bibliografía adjunta.

de hecho, la mejora es tan buena que, si realmente hubiese sido posible construir fibras cuyos perfiles de variación del índice fuesen exactamente iguales al óptimo, la introducción de los sistemas basados en fibras monomodo podría haberse retrasado durante muchos años. El problema es que el mínimo de la dispersión intermodal es tan abrupto (en un primer orden de aproximación, la dispersión intermodal tiende a cero y en un segundo toma el valor dado en la expresión anterior) que es tecnológicamente inviable un proceso de fabricación de fibras con tolerancias lo suficientemente pequeñas que permitan acercarnos a este valor óptimo. Esto se pone de manifiesto en la fig. {2.6}, donde su ha representado también la curva dispersiva de una fibra tecnológicamente viable.

Pasamos, a continuación, a estudiar las fibras ópticas que realmente nos interesan: las monomodo. De entre todas ellas, nos vamos a centrar en las fibras monomodo estándar, que aunque no son las que se suelen instalar en los enlaces de hoy en día, son las únicas que pueden estudiarse analíticamente. Además son las fibras que más se emplearon en la década de los 90 del pasado siglo.

2.3. Ecuaciones de onda

En este apartado queremos determinar los modos de propagación de una fibra monomodo estándar (SMF). Por modo de propagación entendemos una onda monocromática progresiva (esto es, que se propaga) a lo largo del eje longitudinal de la fibra, manteniendo estacionaria la distribución transversal del campo electromanético. Además, la densidad de energía tiende rápidamente a cero a medida que nos alejamos del eje de la fibra.

A partir de las Ecuaciones de Maxwell vamos a deducir las ecuaciones lineales que describen la propagación de ondas monocromáticas a través de fibras SMF, teniendo cuidado de no perder información relevante en esta deducción [5]. Puesto que el medio posee simetría cilíndrica parece, en principio, lógico utilizar este tipo de coordenadas [6].

$$\left.\begin{array}{c} \nabla\times\vec{E} = -\frac{\partial\vec{B}}{\partial t} \\ \nabla\times\vec{H} = \frac{\partial\vec{D}}{\partial t} \\ \vec{B} = \mu_\circ\vec{H};\;\; \vec{D} = \epsilon_\circ\vec{E}+\vec{P} \\ \vec{P} = \epsilon_\circ\displaystyle\int_{-\infty}^{t}\chi(\vec{r},t-\tau)\vec{E}(\tau)d\tau \end{array}\right\} \begin{array}{c} \mathcal{F}_{t\to\omega} \\ \rightleftharpoons \\ \mathcal{F}^{-1}_{\omega\to t} \end{array} \left\{\begin{array}{c} \nabla\times\vec{E}_\omega = -j\omega\vec{B}_\omega \\ \nabla\times\vec{H}_\omega = j_\omega\vec{D}_\omega \\ \vec{B}_\omega = \mu_\circ\vec{H}_\omega;\; \vec{D}_\omega = \epsilon_\circ n^2(\omega)\vec{E}_\omega \\ \vec{P}_\omega = \epsilon_\circ\chi(\omega)\vec{E}_\omega;\; \epsilon_r(\omega) = 1+\chi(\omega) = n^2(\omega) \end{array}\right. \quad (2.7)$$

En el dominio temporal, la integral de convolución que relaciona el momento dipolar macroscópico $\vec{P}$ con el campo eléctrico $\vec{E}$ tiene una interpretación clara: el medio material tiene inercia o, si se quiere, memoria y el valor $\vec{P}$ en el instante actual depende no sólo del valor de $\vec{E}$ en este preciso instante, sino también de los valores que tomó en el pasado. Estamos dando por hecho que χ no depende de $\vec{E}$, lo cual es cierto siempre y cuando el

[5] Con esto queremos decir que las ecuaciones deducidas serán equivalentes a las de Maxwell *en lo que se refiere a la propagación*. No servirán, por supuesto, para describir otros fenómenos electromagnéticos distintos que pudieran darse en la fibra.

[6] Más adelante veremos algo que es contraintuitivo: cuando los índices de refracción de núcleo y corteza son muy parecidos, es más conveniente elegir un sistema de coordenadas cartesianas.

campo eléctrico no sea excesivamente intenso, es decir, cuando la fuerza que se ejerce sobre los osciladores atómicos no es lo suficientemente grande como para que la fuerza atractiva que mantiene al electrón orbitando alrededor del núcleo deje de ser lineal (proporcional al desplazamiento) y el oscilador ya no sea armónico. Aclararemos este punto en el último capítulo de este libro, cuando tratemos efectos no lineales en sistemas de comunicaciones por fibra óptica. También supondremos (y esto es rigurosamente cierto, ya que el sílice es un sólido vítreo y, por tanto, anisótropo) que χ es una función y no un tensor [7].

Parece evidente que las ecuaciones de la derecha (en el dominio de Fourier) son matemáticamente más sencillas que las de la izquierda: ω, al contrario que t, no es una variable sino un parámetro y la integral de convolución se ha transformado en el producto de una función conocida (la susceptibilidad dieléctrica del sílice) por el campo eléctrico (en el dominio espectral, por supuesto). Puesto que queremos estudiar fenómenos de propagación tenemos que llegar a una ecuación de ondas. Conseguirlo es sencillo: obtenemos el rotacional de la primera ecuación de la derecha y tenemos en cuenta las demás:

$$\left.\begin{array}{c}\nabla\times\nabla\times\vec{E}_\omega = -j\omega\mu_\circ(\nabla\times\vec{H}_\omega) = \omega^2\mu_\circ\epsilon_\circ n^2(\omega)\vec{E}_\omega = \frac{\omega^2 n^2(\omega)}{c_\circ^2}\vec{E}_\omega = k_\circ^2 n^2(\omega)\vec{E}_\omega \\ \nabla\times\nabla\times\vec{E}_\omega = \nabla(\nabla\cdot\vec{E}_\omega) - \Delta\vec{E}_\omega = -\Delta\vec{E}_\omega\end{array}\right\}\Rightarrow$$

$$\Delta\vec{E}_\omega + k_\circ^2 n^2(\omega)\vec{E}_\omega = 0 \tag{2.8}$$

Operando de manera análoga con la segunda ecuación, llegamos a una expresión similar en términos del campo magnético. Por tanto, nos queda el siguiente sistema:

$$\left.\begin{array}{c}\Delta\vec{E}_\omega + k_\circ^2 n^2(\omega)\vec{E}_\omega = 0 \\ \Delta\vec{H}_\omega + k_\circ^2 n^2(\omega)\vec{H}_\omega = 0 \\ n(\omega) = n_1(\omega) \quad r\leq a; \qquad n(\omega) = n_2(\omega) \quad r\geq a\end{array}\right\} \tag{2.9}$$

Cada uno de los dos campos tiene tres componentes r, θ y z, por lo que, aparentemente, se propagan seis ondas de manera independiente. Evidentemente, esto no es así [8] y lo que ocurre es que este sistema no es equivalente al de partida y contiene menos información (en estas ecuaciones hay grados de libertad no físicos) que aquél. En realidad sólo dos de las ecuaciones de onda son independientes entre sí y la elección de éstas es, hasta cierto punto, arbitraria. Cuando se trabaja en coordenadas cilíndricas, lo más cómodo es tomar como independientes las componentes z de ambos campos. Teniendo en cuenta la expresión del operador laplaciano en coordenadas cilíndricas, nos quedan las ecuaciones:

[7] Hablando de una manera más rigurosa, χ es un tensor diagonal, con todos sus elementos iguales.

[8] Sabemos que si el medio es isótropo, los campos $\vec{E}$ y $\vec{H}$ han de ser localmente ortogonales entre sí y perpendiculares a su vez al vector de onda.

$$\left.\begin{array}{c}\dfrac{\partial^2 E_{z\omega}}{\partial r^2}+\dfrac{1}{r}\dfrac{\partial E_{z\omega}}{\partial r}+\dfrac{1}{r^2}\dfrac{\partial^2 E_{z\omega}}{\partial \theta^2}+\dfrac{\partial^2 E_{z\omega}}{\partial z^2}+k_\circ^2 n^2(\omega)E_{z\omega}=0\\[2ex]\dfrac{\partial^2 H_{z\omega}}{\partial r^2}+\dfrac{1}{r}\dfrac{\partial H_{z\omega}}{\partial r}+\dfrac{1}{r^2}\dfrac{\partial^2 H_{z\omega}}{\partial \theta^2}+\dfrac{\partial^2 H_{z\omega}}{\partial z^2}+k_\circ^2 n^2(\omega)H_{z\omega}=0\\[2ex] n(\omega)=n_1(\omega)\ \ r\leq a;\qquad n(\omega)=n_2(\omega)\ \ r\geq a\end{array}\right\}\qquad(2.10)$$

Supongamos que ya hemos integrado las dos ecuaciones en derivadas parciales anteriores. ¿Cómo podemos calcular el resto de componentes del campo? Para hacerlo, tenemos que volver a usar las ecuaciones de Maxwell y deducir, a partir de éstas, las expresiones de las componentes transversales de los campos eléctrico y magnético. Después de una serie de operaciones algebraicas, se llega a [9]:

$$\left.\begin{array}{c}\vec{E}_{tr}=\dfrac{1}{k_\circ^2 n^2(\omega)-\beta^2(\omega)}[j\omega\mu_\circ\nabla_{tr}\times\vec{H}_z-j\beta(\omega)\nabla_{tr}\vec{E}_z]\\[2ex]\vec{H}_{tr}=\dfrac{1}{k_\circ^2 n^2(\omega)-\beta^2(\omega)}[-j\omega\epsilon_\circ\nabla_{tr}\times\vec{E}_z-j\beta(\omega)\nabla_{tr}\vec{H}_z]\\[2ex]\nabla_{tr}=\vec{u}_r\dfrac{\partial}{\partial r}+\vec{u}_\theta\dfrac{\partial}{r\partial\theta}\end{array}\right\}\qquad(2.11)$$

En estas expresiones aparece una constante, $\beta(\omega)$ que recibe el nombre de constante de propagación del/los modo(s) guiado(s) de la fibra. Para determinar los modos guiados de la fibra, la manera de proceder es la siguiente: En primer lugar debemos integrar las ecuaciones en derivadas parciales (2.10) y determinar las componentes axiales (componentes según el eje z) de los campos eléctrico y magnético. Al hacerlo, aparecen una serie de constantes de integración cuyos valores hay que determinar imponiendo las condiciones de contorno del problema. Como estamos estudiando fenómenos electromagnéticos en el seno de dieléctricos, estas condiciones de contorno son la continuidad de las componentes normales de los vectores inducción magnética y vector desplazamiento y la de las componentes tangenciales de los vectores campo eléctrico y magnético, tal como se muestra en la fig. {2.9}. Esta continuidad debe verificarse en la superficie de separación entre los dos medio dieléctricos, núcleo y corteza. El resto de componentes no tienen por qué ser continuas en esta interfase. Las componentes normales y tangenciales, se calculan a partir de las axiales empleando las relaciones dadas por las ecuaciones (2.11). Al imponer las conciciones de contorno, se verifica que todas las constantes de integración son nulas (solución trivial: por la fibra no se propaga luz alguna) a menos que $\beta(\omega)$ tome un valor determinado, solución de una ecuación algebraica: es la llamada *ecuación de dispersión.*

Ahora planteamos una de las dos ecuaciones en derivadas parciales, considerando que los índices de refracción de núcleo y corteza no depende de la pulsación ω. De esta manera vamos a determinar exclusivamente el término de la dispersión cromática que depende del guiado de la luz: la llamada *dispersión de guía.* Por tanto, lo que nos queda es:

[9] A partir de ahora dejaremos de utilzar el subíndice ω para simplificar la notación.

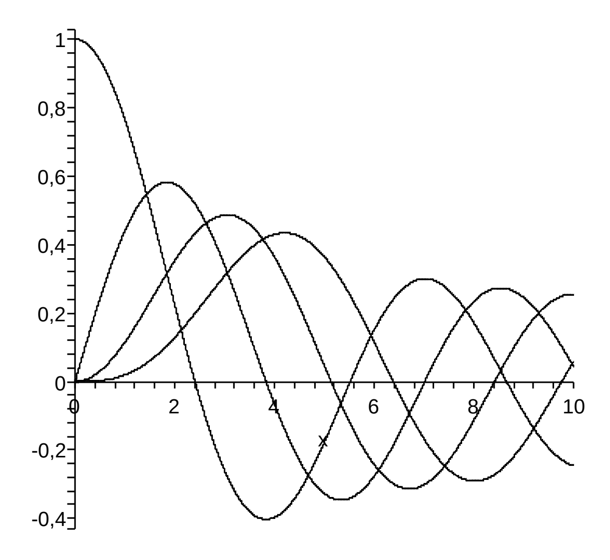


Figura 2.7: Funciones de Bessel de primera especie J_o, J_1, J_2 y J_3.

$$\left.\begin{array}{c}\dfrac{\partial^2 E_z}{\partial r^2}+\dfrac{1}{r}\dfrac{\partial E_z}{\partial r}+\dfrac{1}{r^2}\dfrac{\partial^2 E_z}{\partial \theta^2}+\dfrac{\partial^2 E_z}{\partial z^2}+k_\circ^2 n^2 E_z=0\\[2ex] n=n_1\quad r\leq a;\qquad n=n_2\quad r\geq a\end{array}\right\}\tag{2.12}$$

Puesto que tratamos de encontrar los modos guiados de la fibra, ensayamos en la ecuación anterior soluciones del tipo: $E_z = E_z(r,\theta)e^{-j\beta(\omega)z}$, donde anticipamos que β va a depender de ω aunque consideremos que los índices de refracción de núcleo y corteza toman los valores constantes n_1 y n_2. Esta expresión lo que indica es que las ondas electromagnéticas se propagan a lo largo del eje z, permaneciendo estacionarias en la sección transversal. Sustituyendo, nos queda:

$$\left.\begin{array}{c}\dfrac{\partial^2 E_z}{\partial r^2}+\dfrac{1}{r}\dfrac{\partial E_z}{\partial r}+\dfrac{1}{r^2}\dfrac{\partial^2 E_z}{\partial \theta^2}+[k_\circ^2 n^2 E_z-\beta^2(\omega)]E_z=0\\[2ex] n=n_1\quad r\leq a;\qquad n=n_2\quad r\geq a\end{array}\right\}\tag{2.13}$$

En vez de plantear la ecuación de dispersión por el método más directo (que es puramente analítico), vamos a dar un rodeo y definir dos variables auxiliares, dependientes entre sí, que permiten resolver la ecuación de manera gráfica. Para ello, lo primero que debemos tener en cuenta es que $\beta(\omega)$ es la constante de propagación asociada al modo de pulsación ω. Como parte del campo electromagnético se propaga por el núcleo y parte por la corteza, $\beta(\omega)$ debe tomar valores comprendidos entre las constantes de propagación de una onda plana que viaje por dos medios infinitos de índices n_2 y n_1, esto es $k_\circ n_2 \leq \beta(\omega) \leq k_\circ n_1$. Definimos las siguientes variables positivas:

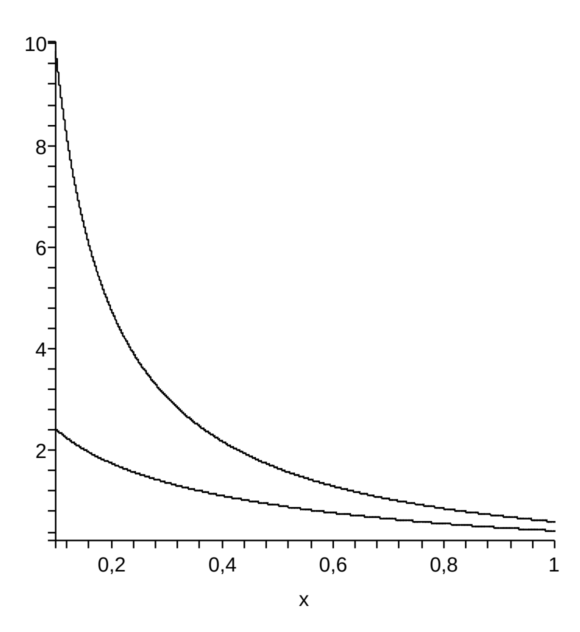


Figura 2.8: Funciones de Bessel de segunda especie K_o y K_1.

$$\left.\begin{array}{l} U^2 := k_\circ^2 n_1^2 - \beta^2(\omega) \\ W^2 := \beta^2(\omega) - k_\circ^2 n_2^2 \end{array}\right\} \Rightarrow U^2 + W^2 = k_\circ^2(n_1^2 - n_2^2) \tag{2.14}$$

Teniendo en cuenta estas definiciones, nos queda:

$$\left.\begin{array}{c} \dfrac{\partial^2 E_z}{\partial r^2} + \dfrac{1}{r}\dfrac{\partial E_z}{\partial r} + \dfrac{1}{r^2}\dfrac{\partial^2 E_z}{\partial \theta^2} + U^2 E_z = 0 \;\; r \le a \\ \dfrac{\partial^2 E_z}{\partial r^2} + \dfrac{1}{r}\dfrac{\partial E_z}{\partial r} + \dfrac{1}{r^2}\dfrac{\partial^2 E_z}{\partial \theta^2} - W^2 E_z = 0 \;\; r \ge a \\ U^2 + W^2 = k_\circ^2(n_1^2 - n_2^2) \end{array}\right\} \tag{2.15}$$

Para integrar las ecuaciones en derivadas parciales (2.15), hay que emplear el método de separación de variable, que consiste en postular soluciones del tipo $E_z = AF_1(r)F_2(\theta)$. Si lo hacemos, llegamos a la conclusión de que la función F_2 verifica la ecuación diferencial ordinaria del oscilador armónico [10]. Además, para que los campos eléctrico y magnético sean unievaluados, las dos soluciones linealmente independientes de esta ecuación son de la forma: $\mathcal{R}\{e^{\pm jl\theta}\}$, donde $l = 0, 1, 2, 3...$ Sustituyendo en las ecuaciones anteriores, obtenemos:

[10] Demuéstrelo como ejercicio.

$$\left.\begin{array}{c} \frac{d^2F_1}{\partial r^2} + \frac{1}{r}\frac{dF_1}{dr} + [U^2 - \frac{l^2}{r^2}]F_1 = 0 \;\; r \leq a \\ \\ \frac{d^2F_1}{\partial r^2} + \frac{1}{r}\frac{dF_1}{dr} - [W^2 - \frac{l^2}{r^2}]F_1 = 0 \;\; r \geq a \\ \\ U^2 + W^2 = k_\circ^2(n_1^2 - n_2^2) \end{array}\right\} \quad (2.16)$$

Estas son las ecuaciones diferenciales de Bessel que aparecen en problemas con simetría cilíndrica. Nos interesan sólo las soluciones con sentido físico, en particular, la función F_1 debe tomar un valor finito cuando $r \to 0$ y debe tender a cero cuando $r \to \infty$. En definitiva, la distribución de campo electromagnético debe tener una energía finita y confinada en las cercanías del núcleo, por lo que resulta:

$$F_1(r) = \left\{ \begin{array}{c} MJ_l(Ur) \;\; r \leq a \\ NK_l(Wr) \;\; r \geq a \end{array} \right. \quad (2.17)$$

Donde J_l es la función de Bessel de primera especia, mientras que K_l es la función modificada de Bessel de segunda especie. El aspecto de estas funciones se muestra en las fig. {2.7} y {2.8}. La soluciones de las ecuaciones en derivadas parciales (2.10) son:

$$\left.\begin{array}{lll} E_z = AJ_l(Ur) & \mathcal{R}[e^{\pm jl\theta}e^{-j\beta(\omega)z}] & r \leq a \\ H_z = BJ_l(Ur) & \mathcal{R}[e^{\pm jl\theta}e^{-j\beta(\omega)z}] & r \leq a \\ E_z = CK_l(Wr) & \mathcal{R}[e^{\pm jl\theta}e^{-j\beta(\omega)z}] & r \geq a \\ H_z = DK_l(Wr) & \mathcal{R}[e^{\pm jl\theta}e^{-j\beta(\omega)z}] & r \geq a \end{array}\right\} \quad (2.18)$$

Donde A, B, C , D y $\beta(\omega)$ son constantes a determinar a partir de las condiciones de contorno.

2.4. Ecuación de dispersión

Para determinar los valores de las constantes de integración, tenemos que imponer las condiciones de contorno en la superficie de separación entre el núcleo y la corteza. A saber, continuidad de las componentes trasversales de los campos $\vec{E}$ y $\vec{H}$ y continuidad de las componentes normales de los campos $\vec{D}$ y $\vec{B}$. Esto es:

$$\left.\begin{array}{lll} E_z)_{r=a^-} = E_z)_{r=a^+} & E_\theta)_{r=a^-} = E_\theta)_{r=a^+} & n_1^2E_r)_{r=a^-} = n_2^2E_r)_{r=a^+} \\ H_z)_{r=a^-} = H_z)_{r=a^+} & H_\theta)_{r=a^-} = H_\theta)_{r=a^+} & H_r)_{r=a^-} = H_r)_{r=a^+} \end{array}\right\} \quad (2.19)$$

Las componentes E_r, E_θ, H_r y H_θ se calculan a partir de E_z y H_z utilizando las relaciones (2.11). Cuando imponemos las condiciones de contorno, que se muestra en la fig.

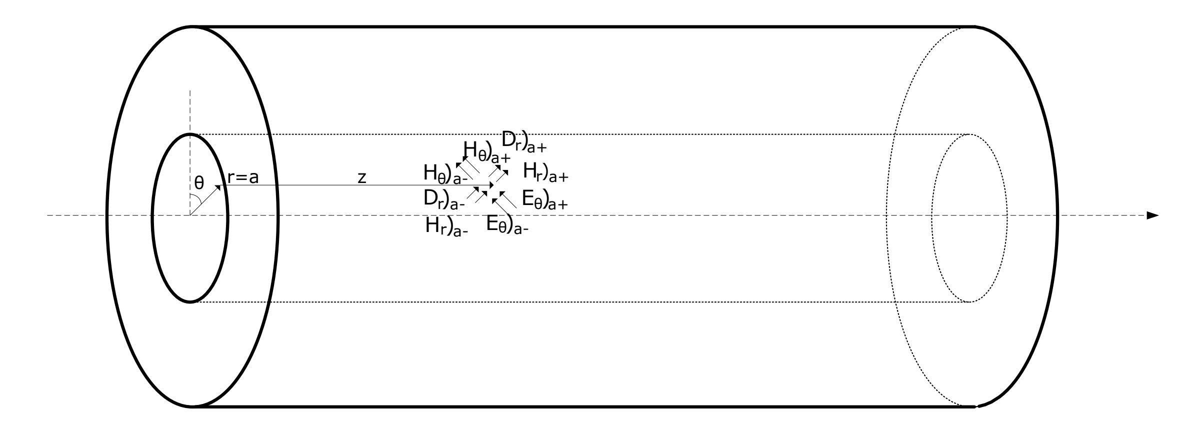


Figura 2.9: Condiciones de contorno en la superficie de separación entre núcleo y corteza.

{2.9}, resulta un sistema de cuatro ecuaciones lineales y homogéneas en A, B, C y D que, evidentemente, siempre tiene la solución trivial ($A = B = C = D = 0$), es decir, que no hay luz propagándose por la fibra. Para que el sistema admita una solución distinta, el determinante del sistema tiene que ser igual a cero (en cuyo caso queda un parámetro libre, que representa la potencia óptica que atraviesa la fibra). Esta condición lleva a que la constante de propagación $\beta(\omega)$ tiene que verificar la siguiente ecuación algebraica [30]:

$$\left[\frac{\dot{J}_l(Ua)}{UJ_l(Ua)} + \frac{\dot{K}_l(Wa)}{WK_l(Wa)}\right]\left[k_o^2 n_1^2 \frac{\dot{J}_l(Ua)}{UJ_l(Ua)} + k_o^2 n_2^2 \frac{\dot{K}_l(Wa)}{WK_l(Wa)}\right] = \beta^2(\omega)\frac{l^2}{a^2}\left[\frac{1}{U^2} + \frac{1}{W^2}\right]^2 \quad (2.20)$$

Ahora vamos a tener en cuenta una hipótesis simplificativa que se cumple siempre en las fibras SMF: la hipótesis de guiado débil [70]–[76]. Esto quiere decir que la diferencia entre los valores de los índices de refracción de núcleo y corteza es muy pequeña: $\Delta \ll 1$ y, por tanto, $n_1 \simeq n_2 \Rightarrow k_o^2 n_2^2 \simeq \beta^2(\omega) \simeq k_o^2 n_1^2$, con lo que (2.20) se reduce a:

$$\frac{\dot{J}_l(Ua)}{aUJ_l(Ua)} + \frac{\dot{K}_l(Wa)}{aWK_l(Wa)} = \pm l\left(\frac{1}{a^2U^2} + \frac{1}{a^2W^2}\right) \quad (2.21)$$

Si resolvemos la ecuación de dispersión de partida (2.20) comprobamos que a medida que $n_2 \to n_1$ existen modos de propagación distintos (distintas distribuiciones del campo electromagnético) cuyas constantes de propagación tienden a tener un mismo valor $\beta(\omega)$, es decir, tienden a degenerar. Tomando ciertas combinaciones lineales estos modos cuasidegenerados [76] que son soluciones de (2.20), se obtienen distribuciones de los campos eléctrico y magnético que, en la sección transversal, tienden a orientarse según dos direcciones ortogonales entre sí a medida que $\Delta \to 0$. Es decir, estas combinaciones de modos están *linealmente polarizadas* (y se denominan modos LP) en el límite $n_2 \to n_1$. Hay que tener en cuenta que en los modos de partida los campos eléctrico y magnético no tienen polarización lineal en la sección transversal. Sólo cuando combinamos linealmente determinados modos

cuasidegenerados (no lo son totalmente ya que $n_1 \neq n_2$) obtenemos modos que sí tienen esta característica. Esta información nos habría permitido plantear las ecuaciones en un sistema de coordenadas cartesianas, que se adapta mejor a la naturaleza de estas soluciones. No se ha hecho así porque la ecuación de dispersión (2.21), al igual que los modos linealmente polarizados (LP), no es más que una aproximación de las soluciones exactas. La aproximación es tanto más válida cuanto menor sea Δ.

Considerando estas combinaciones lineales como los modos propios de la fibra SMF y teniendo en cuenta ciertas propiedades de las funciones de Bessel (propiedades relativas a sus derivadas y de recursividad) se puede demostrar que se llega a la siguiente ecuación de dispersión [76] para los modos LP [11]:

$$\begin{gathered} \frac{aU_{km}J_{k-1}(aU_{km})}{J_k(aU_{km})} = -\frac{aW_{km}K_{k-1}(aW_{km})}{K_k(aW_{km})} \\ k = 0, 1, 2, 3, \ldots \quad U_{km}^2 + W_{km}^2 = k_\circ^2(n_1^2 - n_2^2) \end{gathered} \tag{2.22}$$

Por ahora, el único sentido que tiene el índice m es indicar que, dado una valor de k, la ecuación de dispersión puede tener más de una solución. Para denotar las distintas soluciones, utilizamos este índice. Resulta un poco extraño introducir dos variables distintas U_{km} y W_{km} que depende de $\beta_{km}(\omega)$ y que están relacionadas entre sí por medio de la identidad $U_{km}^2 + W_{km}^2 = k_\circ^2(n_1^2 - n_2^2)$. Planteamos la ecuación de dispersión de esta manera, porque pretendemos resolverla primero gráficamente. La forma de hacerlo sería determinar el rango natural de variación de la variable aU_{km}, representar gráficamente $\frac{aU_{km}J_{k-1}(aU_{km})}{J_k(aU_{km})}$, determinar, después, el rango de variación de $aW_{km} = \sqrt{k_\circ^2(n_1^2 - n_2^2)a^2 - a^2U_{km}^2}$, representar $-\frac{aW_{km}K_{k-1}(aW_{km})}{K_k(aW_{km})}$ y, por último determinar los puntos de intersección de ambas familias de curvas, que nos proporcionarían los valores numéricos de las distintas constantes de propagación de los modos guiados $\beta_{km}(\omega)$.

La constante $V = \sqrt{k_\circ^2(n_1^2 - n_2^2)a^2} = \sqrt{\frac{4\pi^2a^2(n_1^2-n_2^2)}{c_\circ^2}}f$ recibe el nombre de frecuencia adimensional, ya que es un número proporcional a la frecuencia óptica, pero sin dimensiones físicas. Sabemos que $k_\circ n_2 \leq \beta_{km}(\omega) \leq k_\circ n_1$ por tanto, puesto que $a^2U_{km}^2 = [k_\circ^2 n_1^2 - \beta_{km}^2(\omega)]a^2$, deducimos, pues, que aU_{km} toma valores comprendidos en el intervalo $0 \leq aU_{km} \leq V$.

Para una fibra dada, de radio a, índices n_1 y n_2 de valores todos conocido, fijar el valor de la frecuencia adimensional V es equivalente a fijar el valor de la pulsación ω, para la que queremos calcular las constantes de propagación $\beta_{km}(\omega)$ de los modos guiados. Las funciones de Bessel J_k y K_k las evalúan directamente ciertos programas de cálculo numérico, tales como MATLAB. Por consiguiente podemos representar gráficamente:

$$\begin{gathered} f(aU_{km}) := \frac{aU_{km}J_{k-1}(aU_{km})}{J_k(aU_{km})} \\ k = 0, 1, 2, 3, \ldots \;\; 0 \leq aU_{km} \leq V \end{gathered} \tag{2.23}$$

[11] Renombramos el índice l como k ya que en las combinaciones lineales que se hacen para obtener los modos LP, intervienen modos con distintos índices azimutales.

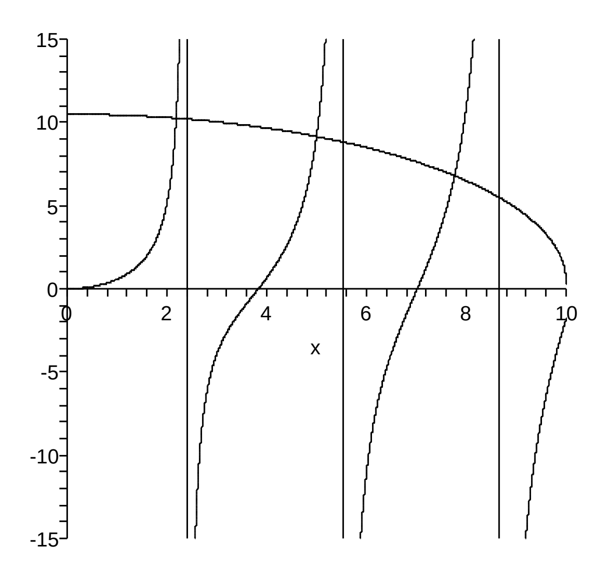


Figura 2.10: Resolución gráfica de la ecuación de dispersión para el caso $k = 0$, con $V = 10$.

Igualmente, puesto que $aW_{km} = \sqrt{V^2 - a^2U_{km}^2}$, podemos representar sobre la gráfica anterior la función:

$$g(aW_{km}) := \frac{\sqrt{V^2 - a^2U_{km}^2}K_{k-1}(\sqrt{V^2 - a^2U_{km}^2})}{K_k(\sqrt{V^2 - a^2U_{km}^2})}$$

$$k = 0, 1, 2, 3, ... \quad 0 \leq aU_{km} \leq V \tag{2.24}$$

Para determinar las constantes de propagación de todos los modos guiados hay que calcular primero V y tomar, inicialmente $k = 0$. Para este valor de k se determinan las distintas soluciones de la ecuación de dispersión y las enumeramos con el índice m. A continuación se toma el valor $k = k + 1$ para el que habrá un número de soluciones igual o inferior al anterior. Eventualmente, llegaremos a un valor máximo de k, k_{max} para el que no existirá intersección alguna entre las dos familias de curvas y habremos determinados las constantes de todos los modos guiados. A modo de ejemplo, se muestra en la fig. {2.10} las tres soluciones que hay para $V = 10$ y $k = 0$.

Se puede comprobar que para que exista un único modo de propagación, que la condiciń que debe cumplirse es $V < 2,405$, en cuyo caso la fibra es monomodo. Si no, tendremos una fibra multimodo a salto de índice. Normalmente en los enlaces que utilizan fibra SMF se trabaja con una V del orden de 2. Este valor es lo suficientemente grande como para que el grado de confinamiento en el núcleo sea importante (más del 70 % de la potencia viaja por el núcleo) y lo suficientemente alejado del valor crítico para el que aparece un nuevo modo guiado.

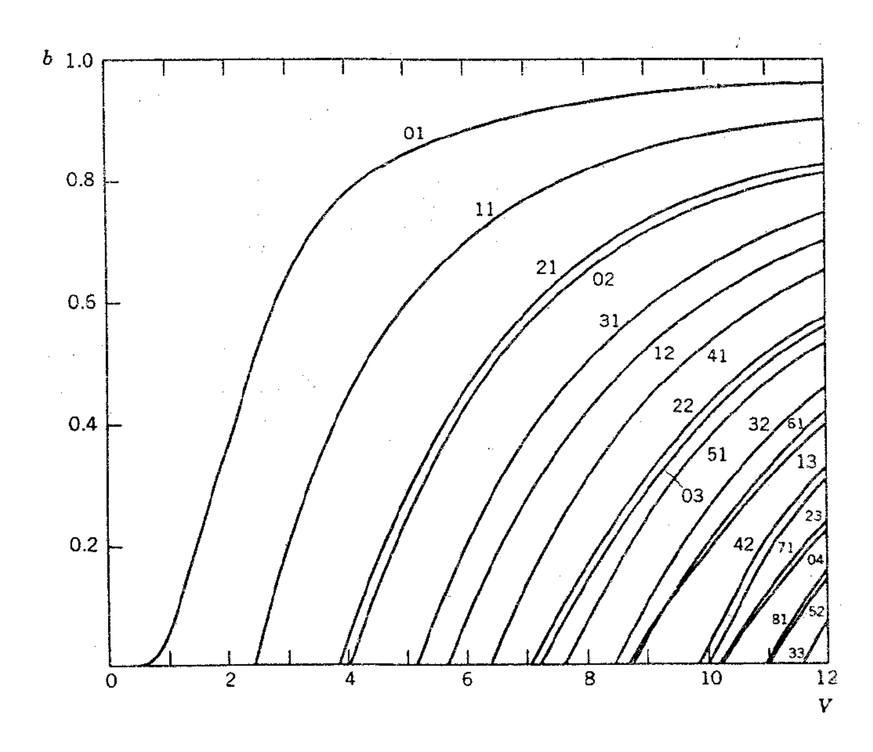


Figura 2.11: Constante de propagación b de los modos LP frente a V. Curva tomada de [38].

2.5. Ecuación de dispersión adimensional

Vamos, a continuación, a plantear la relación de dispersión como una ecuación implícita que relaciona una constante de propagación *adimensional* con la frecuencia adimensional. Para ello, consideremos primero la siguiente expresión:

$$\frac{W^2}{k_\circ^2 n_1^2 - k_\circ^2 n_2^2} = \frac{[\beta(\omega) + k_\circ n_2][\beta(\omega) - k_\circ n_2]}{(k_\circ n_1 + k_\circ n_2)(k_\circ n_1 - k_\circ n_2)} \tag{2.25}$$

Si tenemos ahora en cuenta la condición de guiado débil ($\Delta = \frac{n_1 - n_2}{n_1} \ll 1$), esto implica que $n_1 \simeq n_2$. Además, puesto que $k_\circ n_2 \leq \beta \leq k_\circ n_1$, deducimos que $k_\circ n_2 \simeq \beta \simeq k_\circ n_1$. A la hora de calcular la diferencia entre dos cantidades muy próximas entre sí es muy importante no confundir sus valores, ya que en una expresión en la que aparecen productos de términos, hay una gran diferencia entre que uno de ellos sea pequeño o que sea nulo (un cero en un denominador hace que una expresión deje de tener sentido y en un numerador la anula). Sin embargo, a la hora de calcular una suma, considerar una u otra cantidad es irrelevante. Teniendo esto en cuenta, podemos escribir:

$$b := \frac{W^2}{k_\circ^2 n_1^2 - k_\circ^2 n_2^2} \simeq \frac{\beta(\omega) - k_\circ n_2}{k_\circ n_1 - k_\circ n_2} = \frac{\frac{\beta(\omega)}{k_\circ} - n_2}{n_1 - n_2} = \frac{n_{ef} - n_2}{n_1 - n_2} \tag{2.26}$$

De la expresión anterior deducimos que la constante de propagación adimensional b toma valores comprendidos entre entre 0 y 1 y que representa, en tanto por uno, el índice

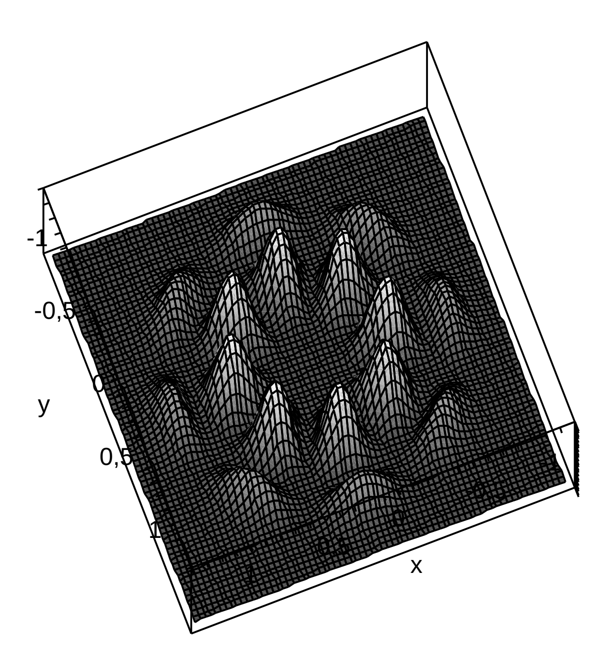


Figura 2.12: Distribución transversal de potencia del modo LP_{42}.

de refracción efectivo del modo considerado. En cualquier caso, estamos ya en disposición de reescribir la ecuación de dispersión en función de una única variable dependiente, b, y de una única variable independiente, V.

$$\left.\begin{array}{r} aW_{km} = V\sqrt{b_{km}} \\ \\ aU_{km} = V\sqrt{1-b_{km}} \end{array}\right\} \Rightarrow \frac{V\sqrt{1-b_{km}}J_{k-1}(V\sqrt{1-b_{km}})}{J_k(V\sqrt{1-b_{km}})} = -\frac{V\sqrt{b_{km}}K_{k-1}(V\sqrt{b_{km}})}{K_k(V\sqrt{b_{km}})} \quad (2.27)$$

El interés de esta ecuación es que es universal y válida para todas las fibras a salto de índice y guiado débil. Sólo hay que resolverla numéricamente una vez (algo que, evidentemente ya está hecho) y representarla gráficamente, tal como se ha hecho en la fig. {2.11}. Los números asociados a cada gráfica son los *índices azimutal y radial* del modo considerado y tienen una clara interpretación física. El primero indica el *números de ceros angulares* de la distribución transversal de la intensidad luminosa del modo en cuestión. Por cero angular entendemos lo siguiente: si trabajamos en coordenadas cilíndricas y existe un valor de la coordenada angular θ para el que la intensidad luminosa es nula para todo valor de r, entonces ese modo de propagación presenta un cero angular. El segundo número nos indica el *número de ceros radiales menos uno.* Por cero radial entendemos que si trabajamos en coordenadas cilíndricas y existe un valor de la coordenada radial r para el que la intensidad luminosa es nula *para todo valor* de θ entonces ese modo de propagación presenta un cero radial. Como ejemplo, mostramos la distribución transversal de potencia del modo LP_{42} en la fig. {2.12}.

Todos los modos salvo el LP_{01} existen (es decir son modos guiados) sólo a partir de un determinado valor de la frecuencia adimensional. La manera de determinar estas frecuencias,

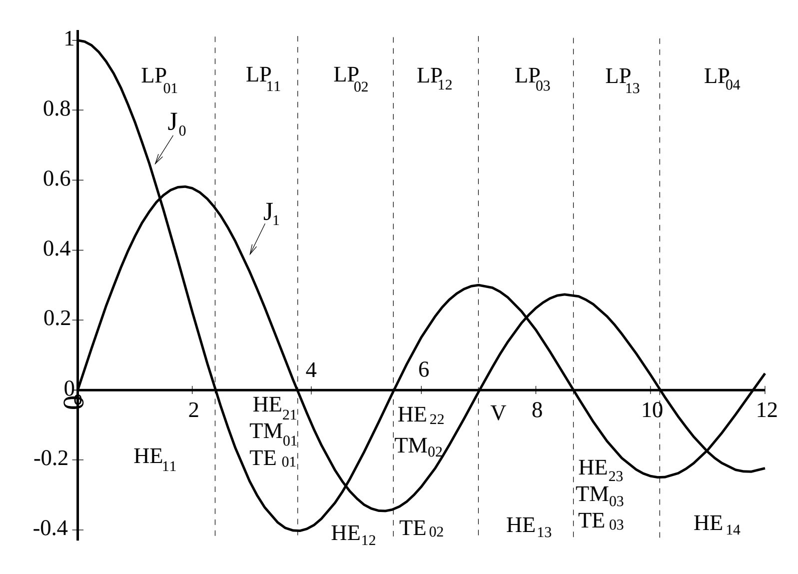


Figura 2.13: Frecuencias de corte de algunos modos LP.

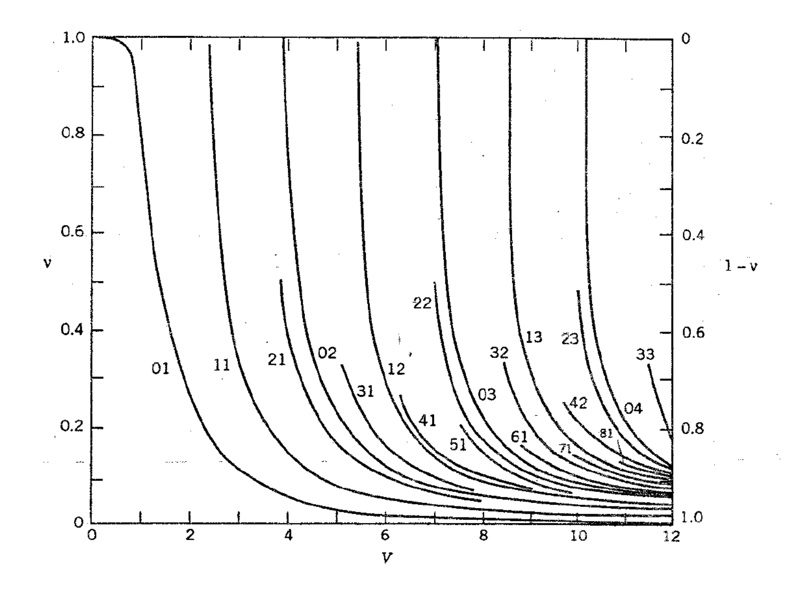


Figura 2.14: Grado de confinamiento de los modos LP_{ml} en función de la frecuencia adimensional V. Curva tomada de [38].

llamadas *frecuencias de corte*, es muy sencilla. La frecuencia de corte del modo LP_{ml} es igual al valor del l-ésimo cero de la función de Bessel J_{m-1}. En la fig. {2.13} se representan algunas funciones de Bessel y se indican qué modos pasan del corte a ser guiados y a qué frecuencias adimensionales ocurre esto.

Consideremos una fibra a salto de índice y guiado débil y un determinado modo LP_{ml} guiado. Evidentemente, cuanto mayor sea la diferencia entre la frecuencia de funcionamiento y la frecuencia de corte de dicho modo, mayor será el grado de confinamiento de la potencia óptica dentro del núcleo de la fibra. Se pueden determinar unas familias de curvas parametrizadas con respecto a los índices m y l , que expresan el porcentaje, en tanto por uno, de la potencia que viaja por la corteza de la fibra con respecto al valor de la frecuencia adimensional V [70]. Estas curvas se representan en la fig. {2.14}.

2.6. Problemas

Problema 3: Considere un enlace de comunicaciones ópticas, en el que la longitud de onda de trabajo es $\lambda_\circ = 1550nm$. El radio de la fibra es $a = 3,4\mu$m y los índices de refracción de núcleo y corteza son $n_1 = 1,45$ y $n_2 = 1,44275$ respectivamente. Compruebe si se trata de un enlace monomodo, determinar las constantes de propagación de los modos guiados y el grado de confinamiento en el núcleo.

Solución: Lo primero que debemos calcular es la frecuencia adimensional y compararla con el valor crítico 2,405 para decidir si se trata de un enlace monomodo o no:

$$\Delta := \frac{n_1^2 - n_2^2}{2n_1^2} \simeq \frac{n_1 - n_2}{n_1} = 0{,}005$$
$$V = k_\circ a\sqrt{n_1^2 - n_2^2} = 2\pi\frac{a}{\lambda_\circ n_1}\sqrt{2\Delta} = 2$$

Por tanto se trata de un enlace monomodo. Si entramos en la fig. {2.11}, deducimos que la constante de propagación adimensional del único modo guiado de la fibra es $b_{01} = 0,4$. Por tanto:

$$\beta_{01} = k_\circ b_{01} = 1621467,2m^{-1}$$
$$n_{01}^{ef} = n_2 + 0,4 \times (n_1 - n_2) = 1,44565$$

Si consultamos la fig. {2.14} llegamos a la conclusión de que el 75 % de la potencia viaja por el núcleo.

Capítulo 3

Atenuación y dispersión en fibras ópticas

Resumen: En este capítulo vamos a estudiar las características de atenuación y dispersión de las fibras ópticas monomodo ya que son las que se usan sistemáticamente en los enlaces troncales. La atenuación está determinada por las propiedades físicas del material con el que se fabrica la fibra, generalmente sílice. La característica dispersiva tiene dos componentes: La dispersión del material (cuyo origen físico ya se ha estudiado en el capítulo 2) y la dispersión de guía que aparece como consecuencia de la diferencia entre los índices de refracción de núcleo y corteza, siendo la dispersión total la suma de estas dos contribuciones, cada una con su propio signo. Existe una tercera componente dispersiva, la llamada dispersión de polarización, que aparece porque el modo LP_{01} es, en realidad, la suma de dos modos degenerados y ortogonales. En presencia de perturbaciones que alteren la simetría cilíndrica de la fibra, los tiempos de propagación asociados a estos dos modos resultan ser distintos y aparece este nuevo término dispersivo. Por último vamos a estudiar cualitativamente las fibras monomodo de dispersión plana (que se emplean en los sistemas DWDM) y las de dispersión desplazada.

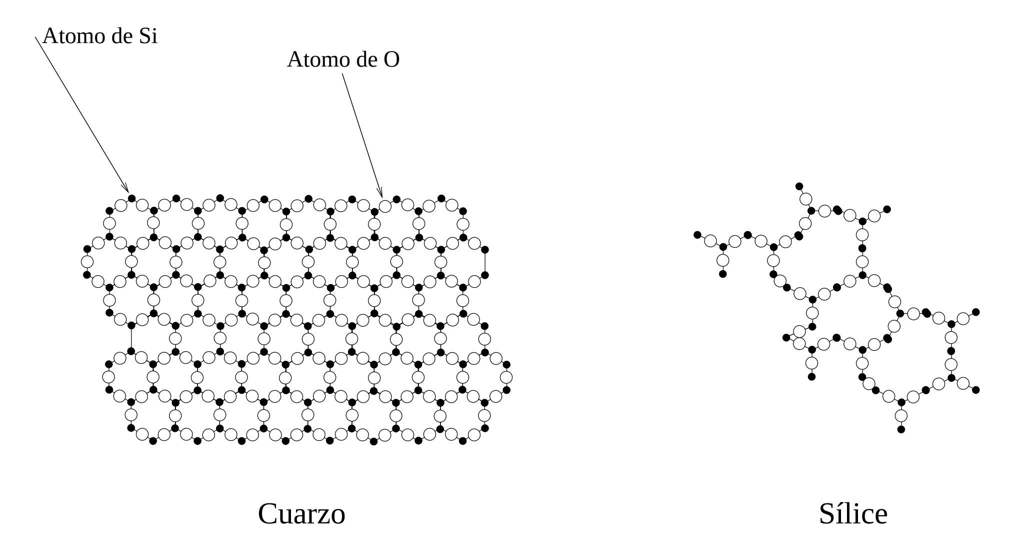


Figura 3.1: Estructuras del cuarzo y del sílice. Encima de cada átomo de silicio hay un átomo de oxígeno que no se ha dibujado, para que la estructura se vea con mayor claridad.

3.1. Introducción

En este capítulo vamos a estudiar los fenómenos de absorción y dispersión en fibras ópticas. Veremos que por ser el sílice un material amorfo (un sólido vítreo) aparece un término de atenuación adicional no estudiado en el capítulo anterior, denominado scattering Rayleigh [18]–[67]. También existe un término dispersivo no relacionado con las propiedades físicas del sílice, sino con el fenómeno de guiado de la luz. Este término recibe el nombre de dispersión de guía y se va a calcular para las fibras monomodo estándar (SMF).

3.2. Microestructura del sílice

El material con el que se fabrica la fibra óptica debe cumplir una serie de requisitos:

- Debe atenuar poco la luz.
- Debe ser un material séctil (fibrable).
- Debe ser ópticamente isótropo.
- Debe ser un material abundante (barato).
- Debe tener una gran resistencia a la tracción.

Todas estas condiciones las cumple el sílice. Tiene un mínimo teórico de atenuación de $0{,}15\frac{dB}{Km}$ a la longitud de onda $1{,}55\mu m$, se pueden fabricar fibras con él fácilmente, es una sólido amorfo y, por consiguiente, isótropo. La materia prima (silicatos) de partida es muy abundante y, por último, es muy resistente a la tracción.

La composición química del sílice, al igual que la del cuarzo, es $Si0_2$. A escala microscópica ambos tienen la misma estructura: un átomo de silicio que se sitúa en el centro

de un tetraedro en cuyos vértices hay átomos de oxígeno. Lo que cambia es la forma en que se combinan estos tetraedros para formar el material [19]. En el cuarzo se ordenan regularmente en las tres direcciones del espacio (se trata de un sólido cristalino), por lo que no todas las direcciones y planos tienen la misma cantidad de átomos. Como consecuencia, las propiedades físicas del cuarzo, y en particular las ópticas, dependan de la dirección en que se midan. Además el cuarzo es un material muy rígido y frágil, por lo que resulta imposible fabricar una fibra óptica con él.

En el sílice la ordenación de los tetraedros es irregular (se trata de un vídreo o sólido amorfo). Por consiguiente no existen direcciones ni planos privelegiadas en el espacio, con lo que el material es isótropo. Además se puede doblar parcialmente sin que se rompa.

Veamos de qué depende que se forme sílice o cuarzo a partir del dióxido de silicio en estado líquido. Si la velocidad de enfriamiento es muy lenta, los tetraedros se ordenan con regularidad en las tres direcciones del espacio de manera que la energía del conjunto sea mínima, formándose así un sólido cristalino llamado cuarzo, cuya estructura se puede ver en la fig. {3.1}. En cambio, si la velocidad de enfriamiento es alta, los tetraedros no alcanzan las posiciones correspondientes al mínimo de energía, produciéndose una estructura desordenada típica de los materiales vítreos y que puede verse en la fig. {3.1}. Este material recibe el nombre de sílice.

Consideremos un volumen de control unos o dos órdenes de magnitud mayor que el volumen de un tetraedro. En el cuarzo, la densidad medida a esta escala es prácticamente independiente de la posición, ya que el material es muy homogéneo. Sin embargo, en el sílice veremos grandes variaciones de la densidad con respecto a la posición del volumen de control, ya que el número de tetraedros contenidos será muy variable. Por tanto, el índice de refracción del cuarzo es uniforme incluso a escalas espaciales muy pequeñas, mientras que el del sílice sufre grandes variaciones si se mide por a una escala suficientemente pequeña.

El volumen de control que, de manera natural, define la luz que se propaga por un medio, es una esfera de radio λ, donde λ es la longitud de onda de la radiación. A medida que λ disminuye, la luz percibe variaciones crecientes del índice de refracción en el sílice. Estas variaciones hacen que exista una cierta probabilidad de que se produzca un proceso de absorción de un fotón y su inmediata reemisión (a la misma longitud de onda) en prácticamente cualquier dirección del espacio. Este fenómeno físico recibe el nombre de scattering Rayleigh [1]. El fotón reemitido sale en la inmensa mayoría de las ocasiones fuera de la fibra, contribuyendo así a aumentar la atenuación. Este término adicional de atenuación es tanto mayor cuanto menor es la longitud de onda de la luz que se propaga por la fibra.

3.3. Curva de atenuación de la fibra óptica

En la fig. {3.2} se muestra la variación del coeficiente de absorción de una fibra en función de la longitud de onda de la luz que se propaga por ella [37]. Vemos que en el infrarrojo cercano (las longitudes de ondas más cortas) el coeficiente de atenuación crece

[1]El scattering Rayleigh es un fenómeno no resonante de interacción luz–materia. Al igual que todos los procesos en los que intervienen estados energéticos virtuales, este mecanismo de interacción es muy débil.

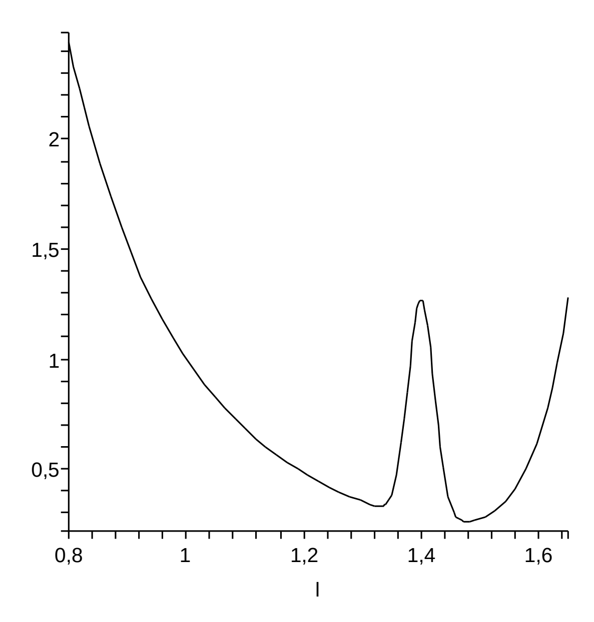


Figura 3.2: Curva de atenuación de una fibra óptica típica. El eje horizontal representa la longitud de onda medida en μm, mientras que el vertical representa la atenuación α medida en $\frac{dB}{Km}$.

rápidamente, a un ritmo que viene dado aproximadamente por $\alpha \simeq \frac{1}{\lambda^4}\frac{dB}{Km}$ (donde λ se mide en micras). En esta zona las pérdidas dominantes son las debidas al scattering Rayleigh. A partir de 1550nm vemos que, de nuevo, la atenuación crece. En este caso el aumento del coeficiente α se debe a la presencia de un pico de absorción en el infrarrojo lejano y en esta región tenemos que $\alpha \simeq 8 \times 10^{11} e^{-\frac{44}{\lambda}} \frac{dB}{Km}$ (aquí también λ se mide en micras). Los estados energéticos responsables de la absorción en esta zona no son electrónicos, sino que están asociados a vibraciones de los enlaces covalentes entre el silicio y el oxígeno.

El pico de absorción que se observa a la longitud de onda de $1,4\mu m$ está provocado por la presencia de iones OH^- en la matriz vítrea del sílice [2]. Si se extreman las precauciones en el proceso de fabricación (humedad ambiente muy baja) de la fibra, este pico no tiene por qué aparecer, aunque, naturalmente, la fibra será más cara. Esta componente de la atenuación es *extrínseca*, ya que no está provocado por las propiedades físicas *intrínsecas* del sílice.

Por último debemos tener en cuenta que esta curva refleja la atenuación de una fibra con simetría cilíndrica perfecta. Cualquier desviación de ésta, como es el curvar la fibra o la presencia de microcurvaturas, provocan un aumento de la atenuación [37]. Las pérdidas serán tanto mayores cuanto menores sean los radios de las curvaturas y microcurvaturas.

El coeficiente de atenuación de una fibra se mide indistintamente en unidades naturales (Km^{-1}), que denotaremos α_n, o bien en decibelios partidos por kilómetro, que denotaremos α_{dB}. Relacionar un coeficiente con el otro es sencillo:

[2] El sílice es un material que tiene una tendencia enorme a hidratarse. De ahí que se utilice como desecante.

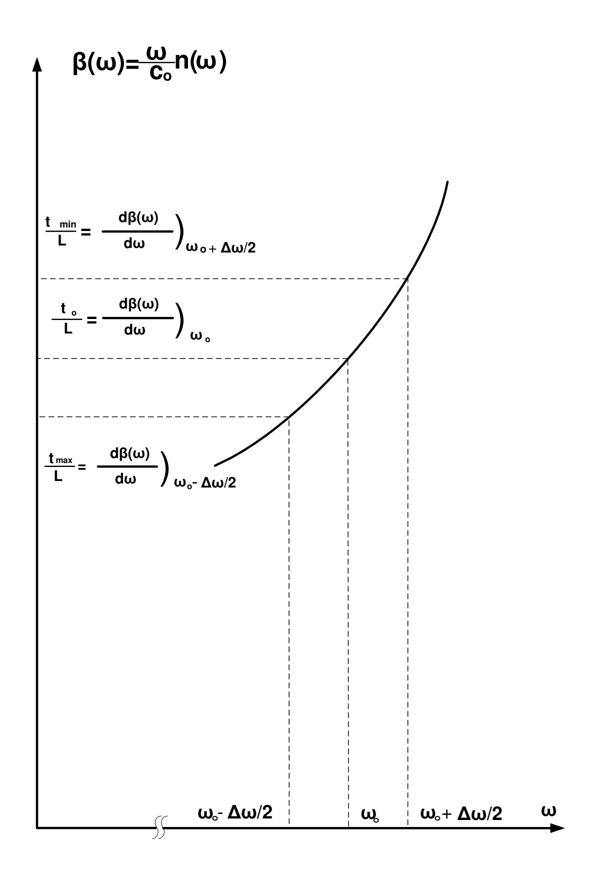


Figura 3.3: Origen físico de la dispersión cromática del material. El caso mostrado es el de un material que presenta una dispersión cromática negativa en $\omega_\circ$.

$$\left.\begin{array}{l} P(z) = P_\circ e^{-\alpha_n z} \\ \alpha_{dB} = \frac{10}{z} Log_{10} \frac{P_\circ}{P_z} \end{array}\right\} \Rightarrow \alpha_{dB} = (10Log_{10}e) \times \alpha_n = 4,343\alpha_n \tag{3.1}$$

3.4. Dispersión cromática del material

Cuando estudiamos la propagación de pulsos luminosos por medios dieléctricos de dimensiones transversales infinitas, vimos que éstos se dispersan como consecuencia de la dependencia no lineal de la constante de propagación, β, con respecto a la pulsación, ω. Dado que para una onda plana $\beta = \frac{\omega}{c_\circ} n$, resulta que un paquete de ondas planas representativo del pulso luminoso sufre dispersión siempre que el índice de refracción depende de ω [3]. En una fibra óptica el índice de refracción del núcleo y el de la corteza dependen de ω de una manera muy parecida y si consideramos que $n_1 \simeq n_2$, podremos calcular una de las dos componentes dispersivas que aparecen en la fibra óptica monomodo: la dispersión cromática del material [41]. Este término depende, pues, de las propiedades físicas del sílice y no del fenómeno de guiado de la luz. Vamos a evaluar de manera aproximada el ensanchamiento temporal asociado.

[3]Esto ocurre siempre en la realidad.

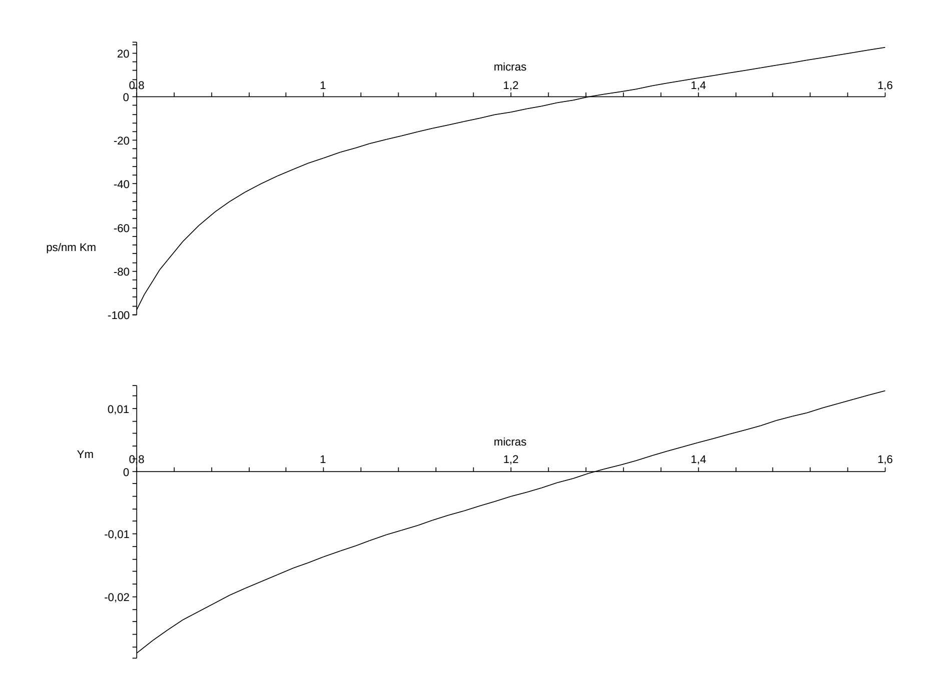


Figura 3.4: Curva de dispersión cromática del sílice con unidades físicas (arriba) y sin ellas (abajo).

La inversa de la velocidad de propagación del pulso (esto es, la inversa de la velocidad de grupo del paquete de ondas que constituyen el pulso) viene dada por:

$$\frac{t_g}{L} = \frac{1}{v_g} = \left(\frac{d\beta}{d\omega}\right)_{\omega_\circ} = \left(\frac{d\beta}{d\lambda}\right)_{\lambda_\circ} \left(\frac{d\lambda}{d\omega}\right)_{\omega_\circ} \tag{3.2}$$

Si ahora tenemos en cuenta que $\beta = \frac{2\pi}{\lambda_\circ} n_\phi(\lambda_\circ)$ y que $\lambda = \frac{2\pi c_\circ}{\omega}$ resulta:

$$\frac{t_g}{L} = \frac{1}{v_g} = \left(\frac{d\beta}{d\omega}\right)_{\omega_\circ} = \frac{1}{c_\circ}\left[n_\phi(\lambda_\circ) - \lambda_\circ \left(\frac{dn_\phi}{d\lambda}\right)_{\lambda_\circ}\right] \Rightarrow n_g = n_\phi(\lambda_\circ) - \lambda_\circ \left(\frac{dn_\phi}{d\lambda}\right)_{\lambda_\circ} \tag{3.3}$$

donde n_ϕ es el índice de refracción de fase (es decir, el que hay que tener en cuenta cuando lo que se propaga por el medio es una onda monocromática) y n_g es el índice de refracción de grupo (el asociado a la propagación de un pulso).

El índice de refracción de grupo, n_g permite calcular la velocidad de un paquete de ondas centrado alrededor de la longitud de onda central del pulso luminoso. En los extremos superior e inferior del espectro del pulso, los índices de grupo toman valores ligeramente distintos a n_g, por lo que las componentes espectrales correspondientes se desplazan con velocidades de grupo ligeramente distintas. Si suponemos que en $\lambda_\circ$ la constante de propagación no presenta un punto de inflexión, podremos estimar el ensanchamiento temporal del

pulso calculando la diferencia entre los tiempos de propagación de grupo en los extremos del intervalo espectral [4]:

$$\left.\begin{array}{l} \frac{t_{g,max}}{L} = \left(\frac{t_g}{L}\right)_{\lambda_\circ+\frac{\Delta\lambda}{2}} \\ \frac{t_{g,min}}{L} = \left(\frac{t_g}{L}\right)_{\lambda_\circ-\frac{\Delta\lambda}{2}} \end{array}\right\} \Rightarrow \frac{c_\circ(t_{g,max}-t_{g,min})}{L} = \frac{c_\circ \tau}{L} = \left(\frac{dn_g}{d\lambda}\right)_{\lambda_\circ} \Delta\lambda =$$

$$= -\lambda_\circ \left(\frac{d^2 n_\phi}{d\lambda^2}\right)_{\lambda_\circ} \Delta\lambda \qquad (3.4)$$

Es costumbre expresar la característica dispersiva del sílice de dos maneras distintas. En la primera el coeficiente dispersivo tiene dimensiones de $\frac{ps}{nmKm}$, mientras que en la segunda el coeficiente es adimensional. Las expresiones que los justifican son:

$$\left.\begin{array}{l} \tau = -\frac{\lambda_\circ}{c_\circ}\left(\frac{d^2 n_\phi}{d\lambda^2}\right)_{\lambda_\circ} \Delta\lambda L \ \ ps \\ \frac{c_\circ \tau}{L} = -\lambda_\circ^2 \left(\frac{d^2 n_\phi}{d\lambda^2}\right)_{\lambda_\circ} \frac{\Delta\lambda}{\lambda_\circ} \end{array}\right\} \Rightarrow \left.\begin{array}{c} D(\lambda_\circ) = -\frac{\lambda_\circ}{c_\circ}\left(\frac{d^2 n_\phi}{d\lambda^2}\right)_{\lambda_\circ} \frac{ps}{nmKm} \\ Y_{\lambda_\circ} = -\lambda_\circ^2 \left(\frac{d^2 n_\phi}{d\lambda^2}\right)_{\lambda_\circ} \end{array}\right\} \qquad (3.5)$$

Las curvas que relacionan los coeficientes D e Y se muestran en la fig. {3.4}.

3.5. Dispersión cromática de guía

La dispersión cromática de guía, como su propio nombre indica, es una componente dispersiva adicional que aparece como consecuencia del fenómeno de guiado de la luz. Si nos centramos en el caso de fibras monomodo estándar (SMF), el guiado de la luz se produce porque existe una zona alrededor del eje de la fibra donde el índice de refracción es más alto (núcleo) que en el resto (corteza). Para calcular la dispersión cromática de guía suponemos que los índices no dependen de la pulsación y que toman los valores constantes n_1 y n_2 en el núcleo y la corteza.

En el capítulo anterior describimos la estructura de una fibra monomodo estándar. Supongamos que el diámetro del núcleo y el salto de índice adimensional son lo sufucientemente pequeños como para que, a la longitud de onda de trabajo, la frecuencia adimensional V sea inferior a 2,405, con lo que existe un sólo modo guiado [5]. $V = 2$ es un valor típico de trabajo para estas fibras, ya que está relativamente alejado de la frecuencia de corte del siguiente modo y el grado de confinamiento de la luz en el núcleo de la fibra es elevado.

Antes de explicar cómo se calcula numéricamente la dispersión cromática de guía a partir de unas curvas adimensionales, vamos a tratar de explicar cualitativamente el motivo

[4]Vamos a suponer que la función es creciente en el intervalo espectral

[5]En realidad son dos modos degenerados que tienen la misma distribución de campos, con la salvedad de que las polarizaciones son ortogonales.

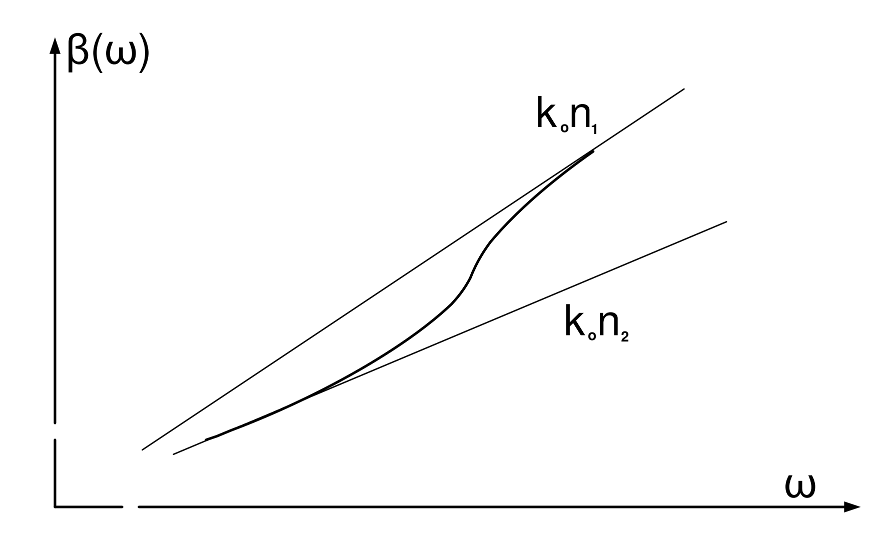


Figura 3.5: Variación de la constante de propagación con respecto a la pulsación.

por el que aparece. Sabemos que la luz que se propaga por la fibra tiene una determinada anchura espectral [6]. En la fig. {3.5} vemos cómo varía el valor de β a medida que aumenta la pulsación: para pulsaciones bajas la mayor parte del campo electromagnético viaja por la corteza, por lo que el valor de β es, prácticamente, el de una onda plana que se propaga por un medio infinito de índice n_2. Para pulsaciones altas ocurre lo contrario: el grado de confinamiento en el núcleo es muy grande y la luz se propaga como lo haría una onda plana por un medio infinito de índice n_1. En la gráfica vemos que el grado de confinamiento de las componentes espectrales de mayor frecuencia es también mayor, por lo que deducimos que los tiempos de propagación de los distintos paquetes de onda cambian con la frecuencia. Como el pulso que se propaga por la fibra tiene una determinada anchura espectral, aparece un término de dispersión cromática asociado al guiado de la luz. Suponiendo que en $\omega_\circ$ no hay un punto de inflexión, los tiempos de propagación máximo y mínimo se producen en los extremos del intervalo y podemos calcular la dispersión cromática de guía calculando la diferencia de las pendientes en los extremos del intervalo:

$$\left.\begin{array}{l} \dfrac{t_{g,med}}{L} = \left(\dfrac{d\beta}{d\omega}\right)_{\omega_\circ} \\ \dfrac{t_{g,max}}{L} = \left(\dfrac{d\beta}{d\omega}\right)_{\omega_\circ+\frac{\Delta\omega}{2}} \\ \dfrac{t_{g,min}}{L} = \left(\dfrac{d\beta}{d\omega}\right)_{\omega_\circ-\frac{\Delta\omega}{2}} \end{array}\right\} \Rightarrow \frac{\tau}{L} = \left(\frac{d\beta}{d\omega}\right)_{\omega_\circ+\frac{\Delta\omega}{2}} - \left(\frac{d\beta}{d\omega}\right)_{\omega_\circ-\frac{\Delta\omega}{2}} = \left(\frac{d^2\beta}{d\omega^2}\right)_{\omega_\circ} \Delta\omega \qquad (3.6)$$

[6]Existen dos casos límite: En el primero la anchura espectral está definida por la propia fuente de luz (esto ocurre en láseres que son poco monocromáticos). En el segundo (láser muy monocromático con modulación externa), es igual a la anchura espectral de los propios datos en banda base.

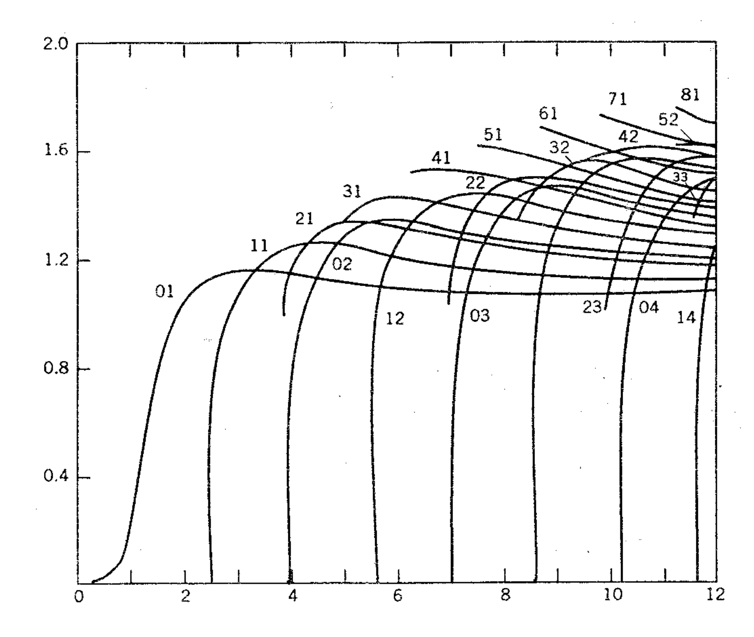


Figura 3.6: Tiempos de propagación adimensionales de los modos guiados de una fibra a salto de índice y guiado débil. Curva tomada de [38].

El tiempo medio de propagación del pulso luminoso por la fibra óptica está definido por el valor de la derivada $\left(\frac{d\beta}{d\omega}\right)_{\omega_o}$, mientras que la dispersión de guía lo está por la segunda derivada $\left(\frac{d^2\beta}{d\omega^2}\right)_{\omega_o}$. Ahora nos gustaría relacionar estas derivadas con las derivadas de la curva universal y adimensional $b = f(V)$. Para hacerlo se parte de las siguientes relaciones:

$$\left.\begin{aligned} b &\simeq \frac{\beta(\omega) - k_o n_2}{k_o n_1 - k_o n_2} = \frac{\beta(\omega) - \dfrac{\omega}{c_o} n_2}{\dfrac{\omega}{c_o} n_1 - \dfrac{\omega}{c_o} n_2} \\ V &= \frac{\omega}{c_o} a \sqrt{n_1^2 - n_2^2} \end{aligned}\right\} \tag{3.7}$$

Si ahora evaluamos la derivada $\left(\frac{d\beta}{d\omega}\right)_{\omega_o}$, teniendo en cuenta las dos igualdades anteriores, llegamos a que el tiempo medio de propagación del pulso luminoso viene dado por la siguiente expresión [70]–[76]:

$$\frac{c_o t}{L} = n_{2g} \left[1 + \Delta \frac{d(Vb)}{dV}\right] \tag{3.8}$$

donde n_{2g} es el índice de refracción de grupo de la corteza y Δ es el salto de índice adimensional de la fibra. Evidentemente, la función $\frac{d(Vb)}{dV}$ es una curva universal válida para todas las fibras de salto de índice y guiado débil. Esta función se obtiene derivando numéricamente la relación de dispersión $b = f(V)$ (obtenida en el capítulo anterior) y está tabulada y representada gráficamente en la fig. {3.6}.

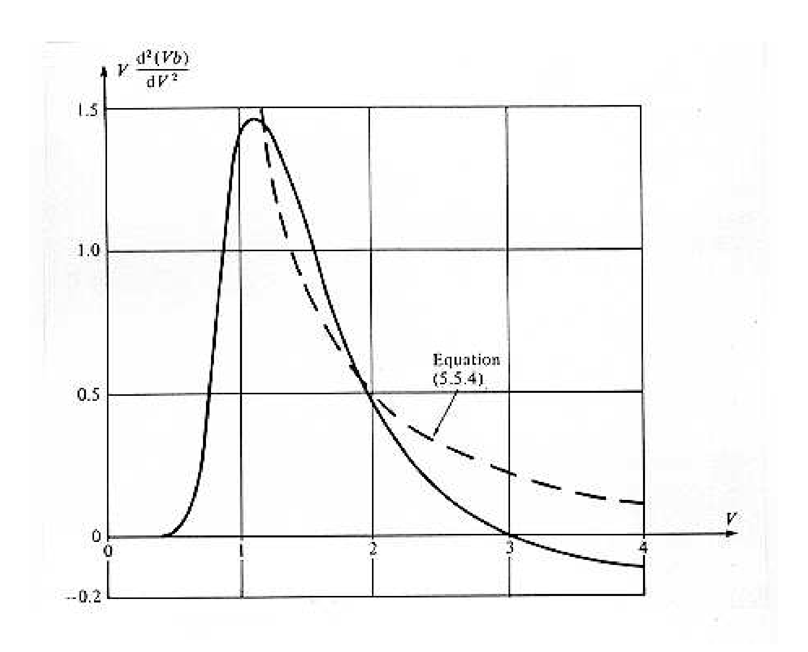


Figura 3.7: Dispersión adimensional de guía de una fibra a salto de índice y guiado débil. En el eje vertical hemos representado $\frac{d(Vb)}{dV}$ y en el horizontal V. Curva tomada de [41].

La forma en la que se deduce el tiempo medio de propagación de un pulso luminoso por una fibra óptica a salto de índice y guiado débil, es derivar la función $\beta(\omega)$ y relacionar esta derivada con las derivadas de la función $b = f(V)$. Resulta, pues, evidente que para calcular la dispersión cromática de guía tenemos que evaluar $\left(\frac{d^2\beta}{d\omega^2}\right)_{\omega_\circ}$ y relacionarla con las derivadas de segundo orden de la función $b = f(V)$. Si hacemos este desarrollo, llegaremos a la siguiente conclusión [70]–[76]:

$$\frac{c_\circ \tau}{L} = \Delta n_{2g} \frac{V d^2(Vb)}{dV^2} \frac{\Delta\omega}{\omega_\circ} = -\Delta n_{2g} \frac{V d^2(Vb)}{dV^2} \frac{\Delta\lambda}{\lambda_\circ} \tag{3.9}$$

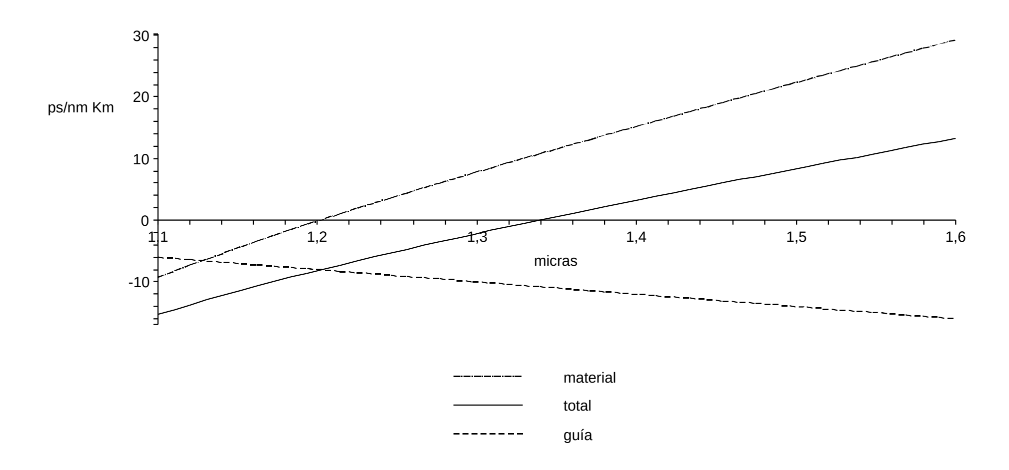
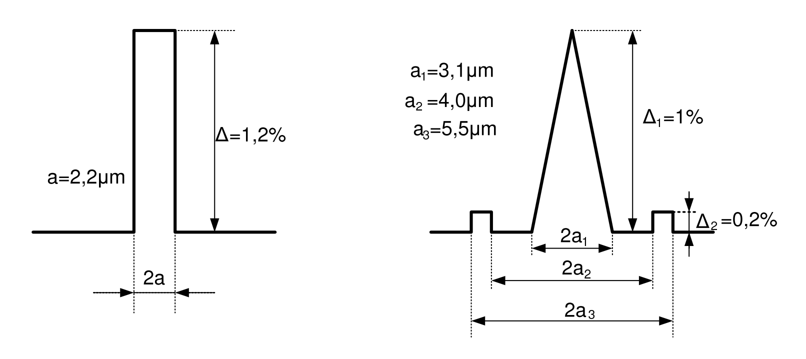


Figura 3.8: Componentes de la dispersión cromática de una fibra SMF.

Hemos representamos la función $\frac{Vd^2(Vb)}{dV^2}$ en la fig. {3.7} para el modo LP_{01} ya que el cálculo de la la dispersión cromática tiene sentido sólo en las fibras monomodo (en las

multimodo la dispersión intermodal es dominante). Vemos que la dispersión cromática de guía es siempre negativa en las fibras SMF, ya que en el intervalo $0 \leq V \leq 2,405$ el signo de $\frac{Vd^2(Vb)}{dV^2}$ no cambia. En la fig. {3.8} representamos las dos componentes de dispersión cromática, así como su suma. *Vemos que en las fibra SMF la dispersión de guía representa sólo una pequeña corrección sobre la del material.* Para fabricar fibras cuyo comportamiento dispersivo sea muy distinto al del material tenemos que usar perfiles que presenten una dispersión cromática de guía muy importante. Los perfiles de variación del índice de refracción necesarios, son muy distintos de los de una fibra SMF. En general estos perfiles no son ni de salto de índice ni de guiado débil.

3.6. Fibras monomodo especiales



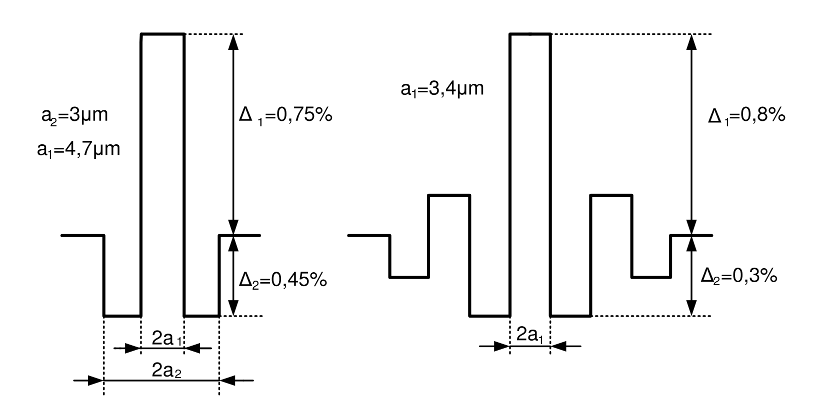


Figura 3.9: Perfiles de variación del índice de algunas fibras especiles.

La fibra óptica que más se utiliza hoy en día no es la monomodo estándar (SMF). En la mayor parte de los tendidos de fibra que se realizan hoy en día, se emplean fibras que tienen la característica dispersiva adecuada para transmitir un número muy elevado de portadoras ópticas multiplexadas en longitud de onda (sistemas DWDM), siendo estas fibras muy distintas de las SMF. El motivo por el que la SMF lleva el adjetivo "estándar"

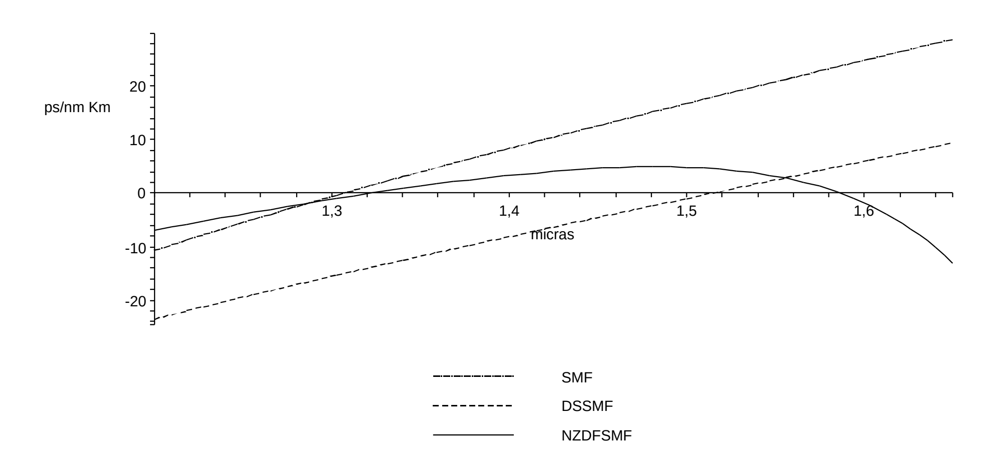


Figura 3.10: Características dispersivas de algunas fibras especiles.

es puramente histórico ya que se utilizaron en la mayoría de los tendidos realizados en la década de los 90 que fue cuando se acuñó este nombre.

Entre las fibras monomodo especiales [59] vamos a destacar dos: la fibra de dispersión desplazada (DSSMF [7]) y la fibra de dispersión plana (NZDFSMF [8]). En la primera de ellas se desplaza el cero de la dispersión cromática total a la longitud de onda de 1550nm, de manera que en tercera ventana haya condiciones óptimas de atenuación y dispersión. Hay que señalar, sin embargo, que el coeficiente de atenuación de este tipo de fibras es superior al de la fibra SMF ya que el nivel de dopado del núcleo de la fibra es sensiblemente más alto (el coeficiente de atenuación de una fibra crece con el nivel de dopado del núcleo [9]). En la segunda, se tiene un coeficiente de dispersión cromática residual no nulo ($\simeq \frac{3ps}{nmKm}$) y *prácticamente constante* en un amplio intervalo de longitudes de onda. Por motivos evidentes este es el tipo de fibra que se utiliza en los sistemas DWDM. Las características dispersivas de estas dos fibras pueden verse en la fig. {3.10}.

La manera de conseguir que la dispersión cromática total sea muy distinta de la del material es utilizar perfiles de variación del índice de refracción cuya dispersión cromática de guía sea importante y la adecuada para conseguir la dispersión cromática total deseada. Los saltos de índices de estas fibras suele ser importantes (para lo cual hay que dopar el núcleo de la fibra con una concentración importante de germanio) y los perfiles son, normalmente, muy distintos del salto de índice, tal como se muestra en la fig. {3.9}. Para estas fibras no existe curvas adimensionales universales y cada fabricante suministra las curvas de dispersión cromática total de sus productos. Las técnicas de diseño de estos perfiles se basan en el cálculo numérico y hay una extensa bibliografía sobre este tema [75].

[7]Del inglés "Dispersion shifted single mode fiber".

[8]Del inglés "Non-zero dispersion flattend single mode fiber".

[9]La manera de conseguir aumentar el índice de refracción de una fibra es sustituir un cierto porcentaje de átomos de silicio por átomos de un dopante, generalmente germanio.

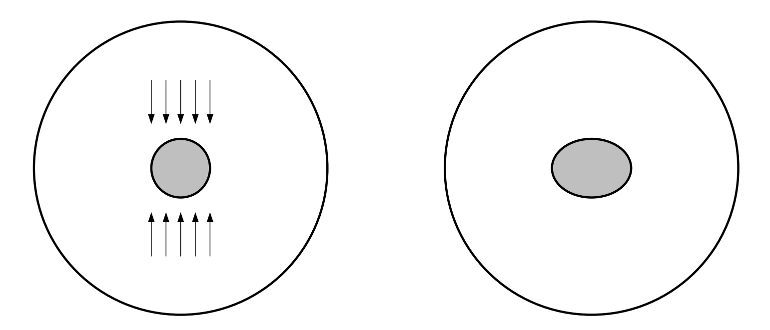


Figura 3.11: Perturbaciones de la simetría de la fibra que producen birrefringencia. Esfuerzos cortantes (izquierda) y elipticidad del núcleo (derecha).

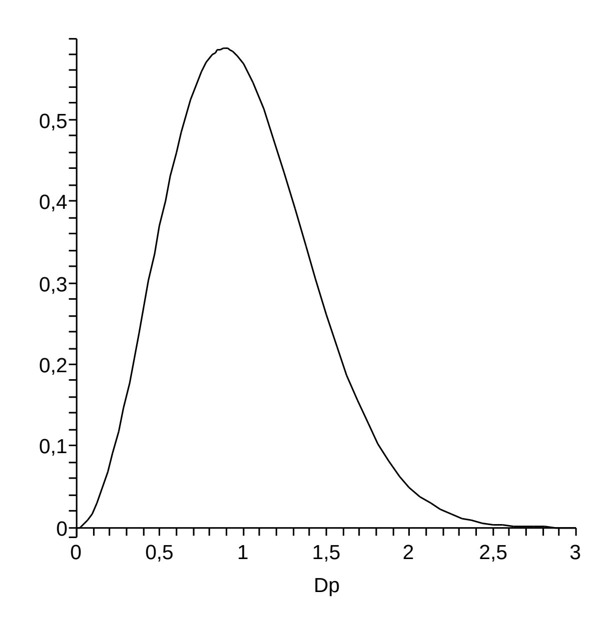


Figura 3.12: Función de densidad de probabilidad de la dispersión de polarización.

3.7. Dispersión de polarización

El modo fundamental de una fibra SMF, el LP_{01}, se compone, en realidad, de dos modos degenerados, esto es, con constantes de propagación que son exactamente iguales y cuyas distribuciones de campo en la sección transversal de la fibra son ortogonales. Esto es así en la medida en la que la fibra presenta una simetría cilíndrica perfecta. En las fibra reales aparecen, sin embargo, perturbaciones de carácter aleatorio que rompen esta simetría y, por tanto, la degeneración de estos dos modos se suprime, convirtiéndose, así en una fibra bimodal. Dos ejemplos de asimetrías que podemos citar son, por un lado, la introducción (accidental) de una cierta elipticidad en el núcleo de la fibra durante su fabricación y, por otro, la aparición de esfuerzos cortantes en el núcleo que inducen una birrefringencia óptica, tal y como se muestra en la fig. {3.11}. Estas perturbaciones cambian de un tramo a otro de la fibra y evolucionan, incluso, a los largo del tiempo.

La birrefringencia inducida en la fibra introduce un nuevo término de dispersión denominado *dispersión de polarización* que es de naturaleza estadística, ya que el lugar y

el instante en el que se producen las asimetrías son aleatorios. El estudio de esta dispersión es muy complicado [40]–[80]–[81]–[82] ya que se basa en los tratamientos con vectores y matrices de Jones, vectores de Stokes y matrices de Müller y la representación sobre la esfera de Poincaré de los estados de polarización de la luz y de los dispositivos modificadores del estado de polarización [10]. Por ello no vamos a desarrollar la teoría, sino que vamos a dar directamente la conclusión a la que se llega.

La dispersión de polarización viene dada por una variable aleatoria, de distribución Maxwelliana. La función de densidad es la siguiente:

$$f(x) = \sqrt{\frac{2}{\pi}} \frac{x^2}{\alpha^3} e^{-\frac{x^2}{2\alpha^2}} \tag{3.10}$$

Es inmediato demostrar que la media de esta variable aleatoria es $\bar{D}_p = \sqrt{\frac{8}{\pi}}\alpha$. Ahora bien, al igual que ocurre en todos los procesos de tipo "random walk" la dispersión de polarización no crece linealmente con la distancia recorrida por los pulsos luminosos sino con su raíz cuadrada, de manera que las unidades de $\bar{D}_p$ son $\frac{ps}{nm\sqrt{Km}}$. En la fig. {3.12} se muestra la función de densidad de esta variable aleatoria. Viendo esta figura resulta claro por qué a la hora de evaluar la dispersión de polarización se toma como dispersión de polarización:

$$\tau_{pol} \simeq 3\bar{D}_p \Delta\lambda \sqrt{L} \tag{3.11}$$

[10]Además es necesaria una cierta familiaridad con la representación de los grupos SU(2) y SO(3).

3.8. Problemas

Problema 4: Se desea diseñar una fibra de dispersión desplazada no nula (NZ-DSF) utilizando el mismo perfil que el de una fibra monomodo estándar. La dispersión residual deseada es de 4ps/nmKm a una longitud de onda de 1550nm. Los índices de refracción de núcleo y corteza son, respectivamente, 1,45 y 1,44275. En caso de existir más de una solución elegir, razonadamente, la más conveniente. Si el espectro de emisión de la fuente es de 1nm, estime la anchura que tendrá un pulso de duración 20ps después de recorrer una distancia de 75 Km

Solución: La dispersión cromática del material viene dada por:

$$\frac{c_\circ \tau_m}{L} = -\lambda^2 \frac{d^2 n}{d\lambda^2} \frac{\Delta\lambda}{\lambda} \Rightarrow$$

$$\frac{\tau_m}{\Delta\lambda L} = -\frac{\lambda}{c_\circ} \frac{d^2 n}{d\lambda^2} = 0,01 \times \frac{1550nm}{3 \times 10^{-7} Km/ps} = 21,5 \frac{ps}{nmKm}$$

Como queremos que al sumarle la dispersión cromática de guía quede 4ps/nmKm deducimos que ésta debe valer -11.5ps/nmKm. Por tanto:

$$\frac{\tau_g}{\Delta\lambda L} = -\frac{\Delta n_2}{\lambda c_\circ} \frac{V d^2(bV)}{dV^2} = -17{,}5 ps/nmKm \Rightarrow$$

$$\frac{V d^2(bV)}{dV^2} = \frac{1}{\Delta n_2} 17{,}5 \frac{ps}{nmKm} \times 1550nm \times 3 \times 10^{-7} \frac{Km}{ps}$$

Como $\Delta \simeq \frac{n_1 - n_2}{n_1} = 5 \times 10^{-3}$ resulta que:

$$\frac{V d^2(bV)}{dV^2} = 17{,}5 \frac{ps}{nmKm} \times 1550nm \times 3 \times 10^{-7} \frac{Km}{ps} \frac{1}{5 \times 10^{-3} \times 1,44275} = 1,128$$

Si consultamos la gráfica $\frac{V d^2(bV)}{dV^2}$ deducimos que V puede tomar dos valores distintos. Tomamos el mayor ya que se corresponde con un radio mayor y, por tanto, con una fibra más fácil de construir.

$$\frac{V d^2(bV)}{dV^2} = 1,128 \Rightarrow V = 1,6$$

$$a = \frac{V\lambda}{2\pi\sqrt{n_1^2 - n_2^2}} = \frac{1,6 \times 1,55 \times 10^{-6}}{\sqrt{1,45^2 - 1,44275^2}} = 2,72\mu m$$

Estimar la anchura del pulso tras reccorrer 75Km es inmediato:

$$\tau_m = T_\circ + D(\frac{ps}{nmKm})\Delta\lambda = 20ps + 4\frac{ps}{nmKm} \times 1nm \times 75Km = 320ps$$

a=$2,72\mu m$
τ_m=320ps

Problema 5: Calcular las características dispersivas de una fibra teniendo en cuenta los siguientes datos: $\lambda_\circ = 1550nm$, $n_1 = 1,45$, $n_2 = 1,44275$, $\Delta\lambda$=2nm y a=$2,44\mu m$. Calcular también el tiempo que tardan los pulsos en recorrer 30 Km.

Solución: Calculemos primero el salto de índice adimensional:

$$\Delta = \frac{n_1^2 - n_2^2}{2n_1} \simeq \frac{n_1 - n_2}{n_1} = 5 \times 10^{-3}$$

A partir de este valor podemos calcular la frecuencia adimensional:

$$V = 2\pi \frac{a}{\lambda_\circ} n_1 \sqrt{2\Delta} = 1,434 < 2,403$$

Por tanto la fibra es monomodo y no hay dispersión intermodal

$$\frac{c_\circ \tau_m}{L} = -\lambda_\circ^2 \frac{d^2 n}{d\lambda^2} \left| \frac{\Delta\lambda}{\lambda_\circ} \right| = 0,01 \times \frac{2}{1550} = 1,3 \times 10^{-5}$$

$$\frac{c_\circ \tau_g}{L} = \Delta n_2 \frac{V d^(V b_)}{dV^2} \left| \frac{\Delta\lambda}{\lambda_\circ} \right| = 10^{-5} \Rightarrow$$

$$\frac{c_\circ \tau_t}{L} = -0,3 \times 10^{-5}$$

Por tanto: $\tau_t = 168ps$

$$\frac{c_\circ t_{prop}}{L} = n_2 \left[1 + \Delta \frac{d(V b_{01})}{dV} \right] \simeq 1,456 \Rightarrow$$

$$t_{prop} = 145,6\mu s$$

$$\boxed{\begin{array}{l} \tau_t = 168ps \\ t_{prop} = 145,6\mu s \end{array}}$$

Capítulo 4

El láser de niveles discretos

Resumen: En este capítulo vamos a estudiar el láser de niveles discretos. Este dispositivo no se emplea en comunicaciones ópticas, pero comparte con el resto de láseres los mismos principios básicos de funcionamiento; además, los fenómenos de de interacción luz–materia que se producen en su medio activo son menos complejos que los de los láseres semiconductores, que son los dispositivos empleados en comunicaciones por fibra óptica. El motivo es que en un láser de niveles discretos los estados energéticos accesibles están cuantizados y no forman bandas de energía, como en los semiconductores.
Para entender el funcionamiento de los láseres tenemos que introducir primero un mecanismo de interacción luz–materia aún no estudiado: la emisión estimulada. Hallando el balance del número neto de fotones añadidos a la radiación electromagnética como la diferencia entre los generados por emisión estimulada menos los robados por absorción, podemos determinar la ganancia de un medio material y, por tanto, las condiciones de emisión de luz láser.

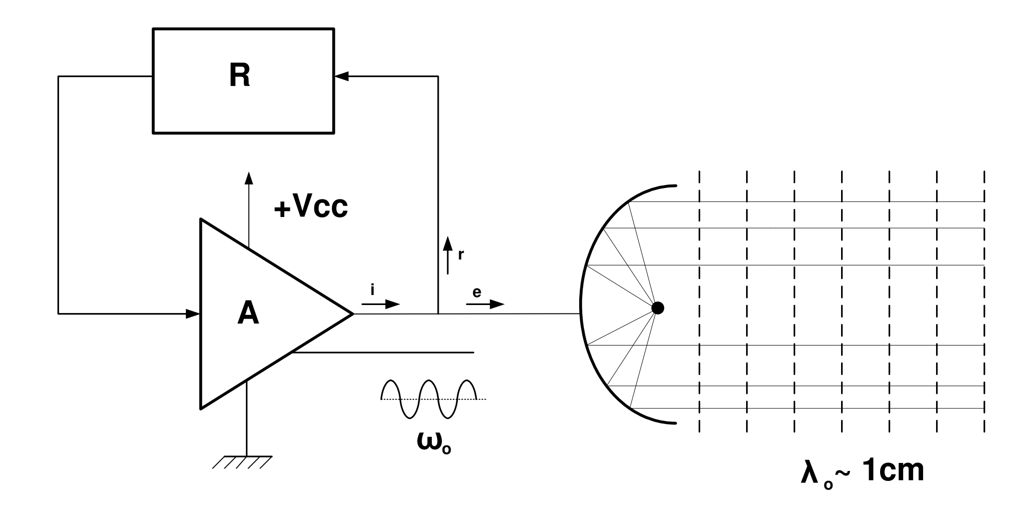


Figura 4.1: Esquema general de un emisor de ondas de radiofrecuencia.

4.1. Introducción

En capítulos posteriores, estudiaremos el transmisor que se utiliza normalmente en comunicaciones ópticas (sobre todo en enlaces en los que se transmiten tasas binarias elevadas a grandes distancias): el láser a semiconductor. Sin embargo existe un tipo de láser, que aunque no se emplea en comunicaciones ópticas, es relativamente sencillo de estudiar: *el láser de niveles discretos*. Constan de los mismos elementos básicos que un láser semiconductor y presenta la ventaja de que el medio material amplificador tiene unos estados energéticos accesibles de estructura mucho más simple. Por este motivo, como paso previo al estudio de los láseres que realmente nos interesan, abordamos el funcionamiento de los láseres de niveles discretos.

Un láser es, básicamente, un oscilador o generador de ondas electromagnéticas y su estructura y elementos básicos son los mismos que los de un oscilador de radiofrecuencia, con la particularidad de que la longitud de onda de emisión es bastante corta, del orden de una micra. En la fig. {4.1} se muestran los elementos básicos de una emisora de radiofrecuencia, que son:

1. **Un amplificador**, encargado de reproducir a la salida con mayor amplitud cualquier radiación electromagnética que haya a su entrada, siempre que caiga dentro de su ancho de banda de funcionamiento.

2. **Una fuente de alimentación** que cede al medio amplificador la potencia necasaria para realizar esta función.

3. **Una red o bucle de realimentación** que fija la pulsación de la radiación electromagnética emitida por el oscilador.

4. **Una antena** que emita la radiación generada al espacio radioeléctrico.

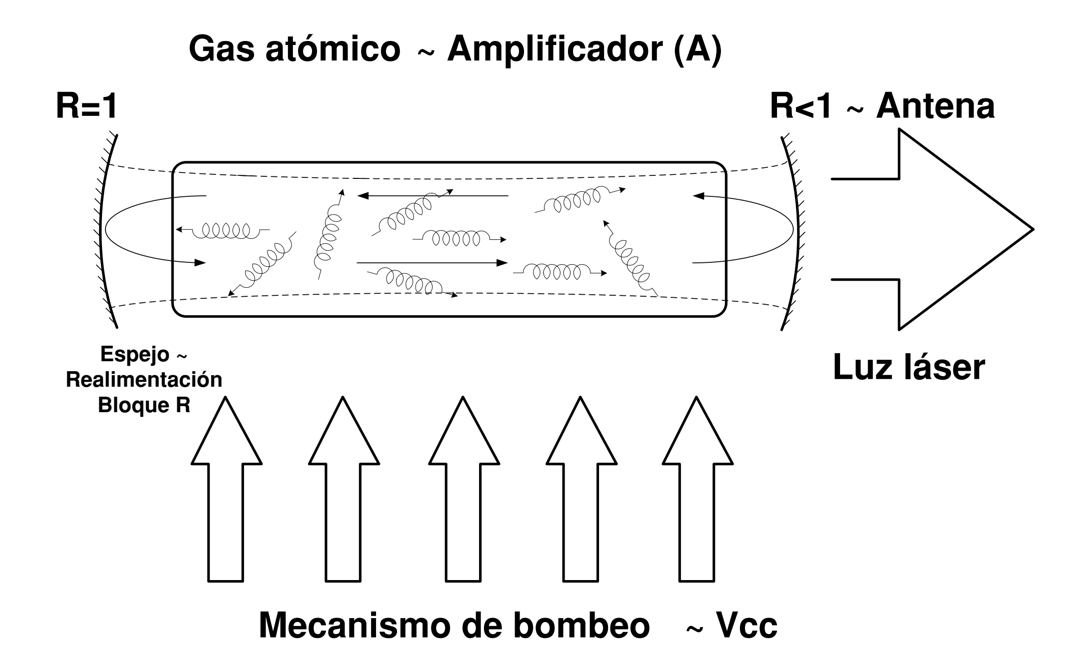


Figura 4.2: Estructura de un láser y su analogía con la de un emisor de ondas RF.

El medio material de un láser de niveles discretos es, en muchas ocasiones, un gas atómico como el descrito en el capítulo 1. Sobre este medio se actúa externamente suministrando una cierta potencia (este proceso se conoce con el nombre de *bombeo*) para conseguir que el gas tenga un coeficiente de *ganancia* de manera que se amplifique la luz que se propaga en su seno (siempre que caiga dentro de un determinado intervalo espectral). El gas está metido en un tubo cilíndrico hermético, limitado en sus extremos por dos espejos que tienen un coeficiente de reflexión en potencia alto. El volumen definido por el tubo y los dos espejos se denomina *cavidad resonante* y es el equivalente a la red de realimentación de un emisor de ondas de RF y determina la longitud de onda de emisión $\lambda_\circ$. Para que se produzca la emisión de luz, uno de los espejos tiene un coeficiente de reflexión inferior a la unidad, de forma que deja pasar al exterior parte de la radiación electromagnética incidente y es el equivalente a la antena emisora del oscilador de RF. En la fig. {4.2} se muestra la estructura de un láser de niveles discretos y se compara con la de una emisora de radiofrecuencias (RF).

Las características de la luz emitida por un láser son las mismas que las de la radiación electromagnética generada por una emisora de RF, a saber:

1. **Luz espectralmente pura.**

2. **Luz perteneciente a un único modo espacial.** Esto quiere decir que la luz se propaga en una dirección específica del espacio [1].

3. **Luz polarizada.** Al igual que la radiación generada por una emisora de radio, el estado de polarización es fijo.

[1] Esto es debido a que la cavidad resonante es muy alargada. Si la cavidad resonante fuese, por ejemplo, esférica, el láser emitiría ondas esféricas.

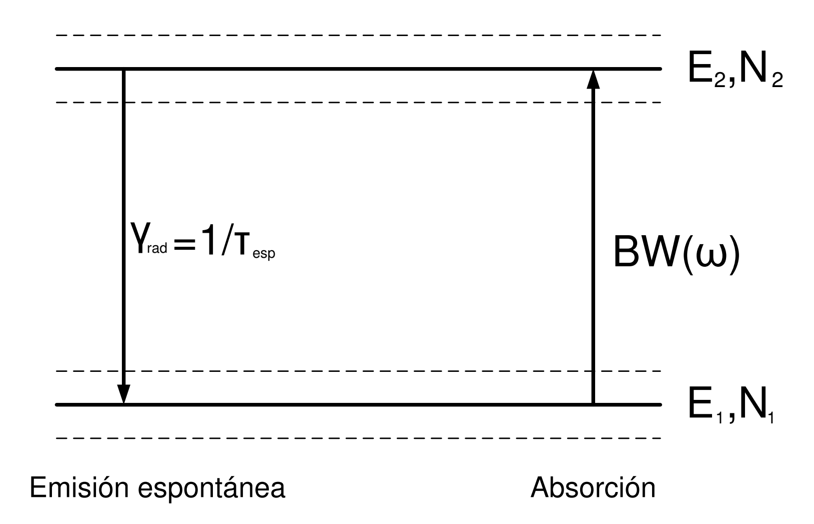


Figura 4.3: Diagrama de estados energéticos accesibles de un átomo y procesos clásicos de interacción luz–materia.

4.2. Coeficientes de Einstein

En el primer capítulo describimos de una manera simplificada los mecanismos de interacción entre la luz y un medio material particular: un gas atómico. Vimos que existían dos mecanismos de interacción luz–materia: El primero, la absorción, se manifiesta claramente cuando se propaga por el medio una radiación electromagnética cuya pulsación ω toma un valor cercano a la pulsación característica del átomo $\omega_\circ$. La absorción se produce porque los átomos (u osciladores mecánicos en el modelo de Lorentz) roban energía electromagnética a la radiación que se propaga, aumentando su energía mecánica (potencial y cinética) de oscilación. El segundo proceso es la emisión espontánea, que aparece como consecuencia del término de fricción γ del oscilador mecánico y consiste en la emisión por parte de los átomos de fotones con pulsación, fase, estado de polarización y dirección de propagación relativamente aleatorios.

Existe un tercer mecanismo de interacción luz–materia sin el que sería imposible el funcionamiento del láser: la *emisión estimulada*. El objetivo de esta sección es incluirlo en nuestro modelo de interacción y describir los anteriores procesos (absorción y emisión espontánea) teniendo en cuenta que, en realidad, los estados energéticos accesibles para los átomos están discretizados, tal y como se muestra en la fig. {4.3}. Aunque siempre hay más de dos estados energéticos accesibles, no vamos a considerar los demás con objeto de simplificar el estudio.

El modelo mecánico introducido en el capítulo 1 predice resultados cualitativamente ciertos para medios materiales dieléctricos bajo condiciones relativamente próximas al equilibrio térmico pero, además, podemos modificarlo para incluir en él la emisión estimulada.

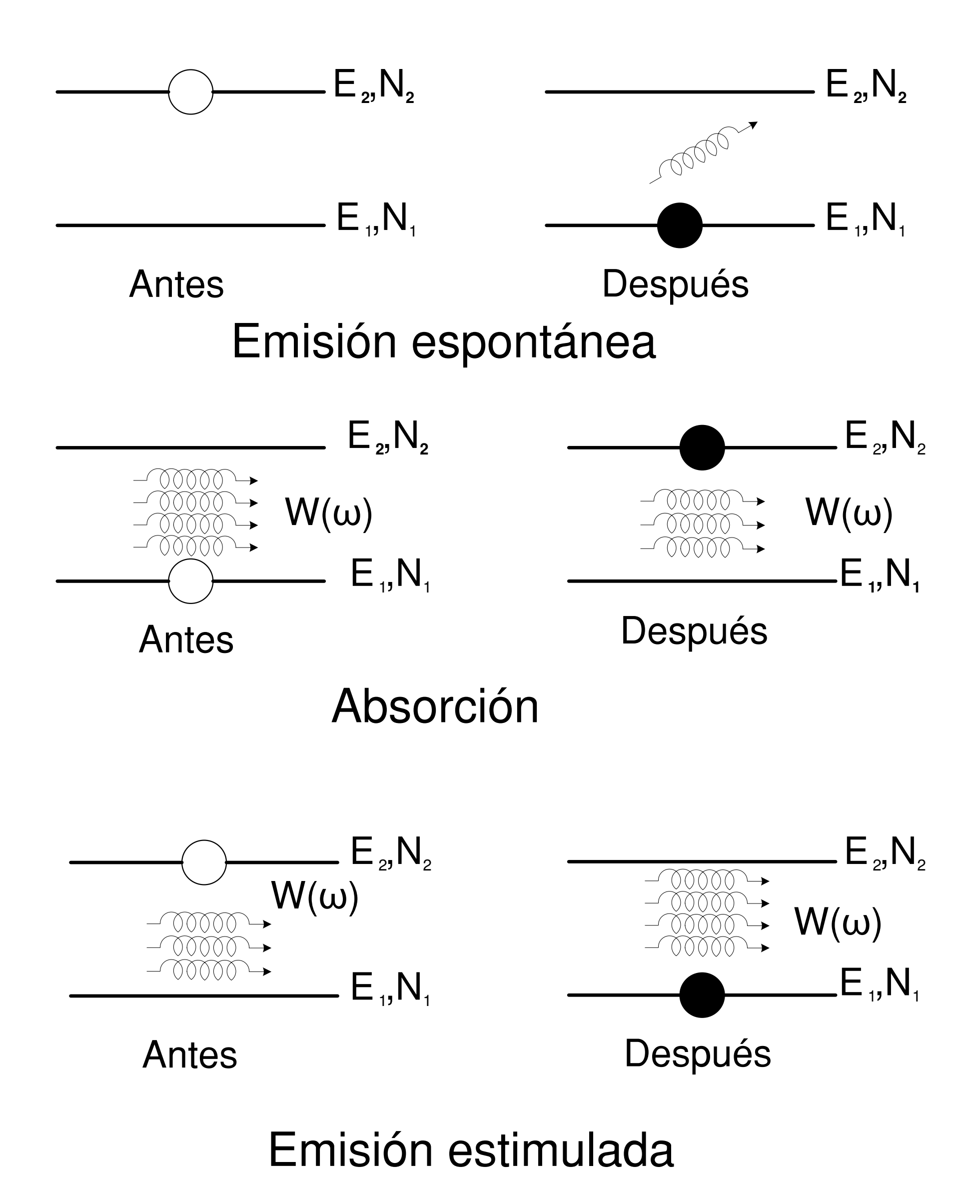


Figura 4.4: Procesos elementales de interacción luz–materia.

La relación que existe entre la pulsación característica $\omega_\circ$ y los parámetros de la fig. {4.3} es $\omega_\circ = \frac{E_2 - E_1}{\hbar}$. Sabemos, además, que los átomos pueden absorber y emitir radiación espontánea de pulsaciones ligeramente distintas a $\omega_\circ$, lo que quiere decir que los estados energéticos E_1 y E_2 tienen una cierta incertidumbre [2] en cuanto a su valor, tal como se muestra en la figura. Por tanto, los procesos de interacción luz–materia no tienen por qué producirse exactamente a la pulsación $\omega_\circ$, sino que pueden producirse en un intervalo estrecho de pulsaciones alrededor de este valor. La función de densidad de probabilidad de esta variable aleatoria (pulsación o frecuencia a la que se produce un proceso de interacción luz–materia) recibe el nombre de *función de línea.*

La primera descripción cuántica de los procesos de interacción luz–materia la dió Einstein en 1915 con un modelo fenomenológico [67]–[98] y en ella, además de los procesos de

[2] Si ninguno de los dos estados es el de mínima energía, los tiempos medios de permanencia de un átomo en ellos son finitos y la segunda relación de incertidumbre de Heissenberg afirma que los valores energéticos no están bien definidos.

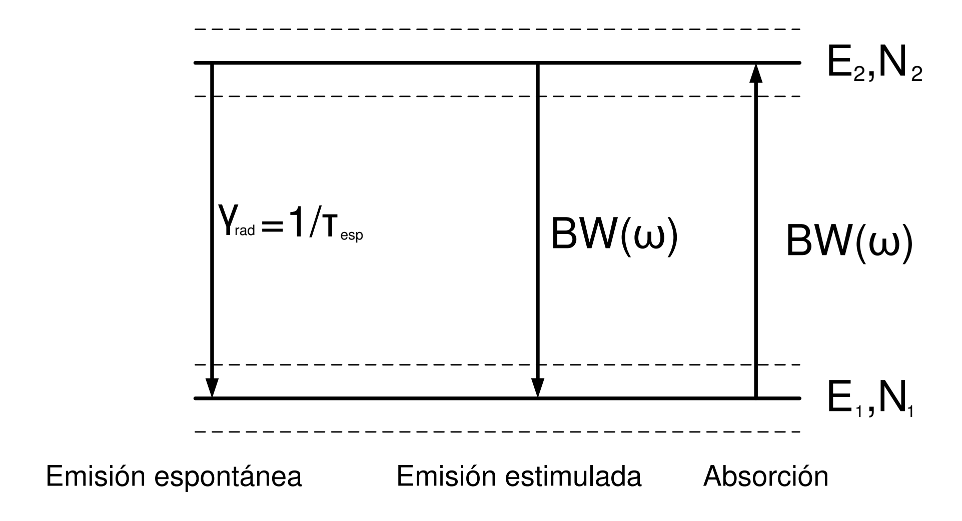


Figura 4.5: Diagrama completo de los procesos de interacción luz–materia.

absorción y emisión espontánea, introdujo un nuevo mecanismo de interacción: la emisión estimulada. Las ecuaciones que describen estos fenómenos son las siguientes [3]:

$$\left.\begin{array}{l} \left(-\dfrac{\partial N_2}{\partial t}\right)_{esp} = \dfrac{N_2}{\tau_{esp}} l(\omega) d\omega \\ \left(\dfrac{\partial N_2}{\partial t}\right)_{abs} = BW(\omega) l(\omega) N_1 d\omega \\ \left(-\dfrac{\partial N_2}{\partial t}\right)_{est} = BW(\omega) l(\omega) N_2 d\omega \end{array}\right\} \quad (4.1)$$

En estas ecuaciones N_2 representa el número de átomos que, por unidad de volumen, ocupan el estado energético "2". Este número puede aumentar gracias a la absorción y disminuir por culpa de la emisión espontánea y por el aún no descrito mecanismo de emisión estimulada. El término $W(\omega)$ representa la densidad espectral de energía electromagnética por unidad de volumen, mientras que N_1 indica el número de átomos que, por unidad de volumen, ocupan el estado energético "1". Este número puede crecer gracias a los mecanismos de emisión espontánea y emisión estimulada y disminuir por culpa del mecanismo de absorción. Los parámetros B y $\gamma_{rad} = \frac{1}{\tau_{esp}}$ son los coeficiente fenomenológicos de Einstein, cuyo valores teóricos sólo pueden calcularse acudiendo a mecánica cuántica [4].

En la fig. {4.4} se representan los tres procesos arriba mencionados [5]. En el primero,

[3]Suponemos que el grado de degeneración para los estados "1" y "2" es el mismo.

[4]Si empleamos la "regla de oro" de Dirac (injustamente conocida como "regla de oro de Fermi") [28] los cálculos se pueden hacer en el marco de una teoría semiclásica. Si no, hay que acudir a la electrodinámica cuántica [22]–[42]–[50]–[67].

[5]En realidad no es lícito representar los fotones como partículas localizadas ni considerar un único átomo.

la emisión espontánea, se produce el paso de un átomo del estado energético superior al inferior y la emisión de un fotón de energía cercana a $E_2 - E_1$, propagándose en cualquier dirección del espacio, con cualquier estado de polarización y con cualquier referencia de fase. Hay que destacar el hecho de que este proceso ocurre *independientemente* de la presencia o no de radiación electromagnética en el medio [6] y de ahí el sentido del adjetivo espontáneo. Los otros dos procesos están *estimulados* por la presencia de un campo electromagnético de pulsación cercana a $\omega_\circ = \frac{E_2 - E_1}{\hbar}$ y, de alguna manera, son el inverso el uno del otro. El segundo proceso, la absorción, supone el paso de un átomo del estado energético inferior al superior, tomando la energía necesaria de la radiación electromagnética que se propaga por el medio (es decir, se roba un fotón y el átomo pasa al estado de mayor energía). Por último, la emisión estimulada consiste en el paso de un átomo del estado superior al inferior, estimulado por la presencia de radiación electromagnética de pulsación cercana a $\omega_\circ = \frac{E_2 - E_1}{\hbar}$. La energía perdida se emplea en aumentar en uno el flujo de fotones que atraviesa el medio atómico.

El número de fotones producidos o robados por un proceso de interacción luz–materia, en un intervalo espectral $(\omega, \omega + d\omega)$ se escribe como el producto de una serie de términos:

1. Un coeficiente de proporcionalidad, que es γ_{rad} para el proceso espontáneo y B para los procesos estimulados.

2. La densidad de átomos que corresponda: N_1 para la absorción y N_2 para los procesos de emisión.

3. La densidad espectral de energía $W(\omega)$ (por unidad de volumen) para los procesos estimulados.

4. La probabilidad de que se produzca un proceso de interacción a esa pulsación $l(\omega)d\omega$.

El significado de la función de línea parece claro: puesto que existe una cierta banda de incertidumbre en cuanto a la posición de los estados energéticos E_1 y E_2, pueden producirse procesos de interacción en un entorno de pulsaciones alrededor de $\omega_\circ$ [7]. Como la interacción tiene que producirse a alguna pulsación, queda claro que $l(\omega)$ verifica que $\int_0^\infty l(\omega)d\omega = 1$ como cualquier función de densidad de probabilidad.

Ya estamos en disposición de representar e interpretar el diagrama completo de los procesos de interacción luz–materia. Éste se muestra en la fig. {4.5}. Einstein, a partir de consideraciones termodinámicas [98], dedujo la existencia de la emisión estimulada y determinó la relación que existe entre el coeficiente asociado a los procesos estimulados (B)

[6] En algunos libros de electrodinámica cuántica se afirma que la emisión espontánea está, en realidad, estimulada por las fluctuaciones de la energía de vacío del campo electromagnético y que ésta, junto con el efecto Casimir y la renormalización de masa y carga, son las manifestaciones físicas más evidentes del vacío electromagnético. Interpretaciones más profundas y recientes de la emisión espontánea [73] matizan y corrigen esta interpretación.

[7] En realidad hay otros procesos que contribuyen a ensanchar aún más la función de densidad [92]. Los hay que afectan por igual a toda la población atómica, en cuyo caso se tiene una función de línea homogenea (sería el caso en el que las colisiones atómicas elásticas fuesen el fenómeno dominante). En otras ocasiones los procesos de ensanchamiento espectral "clasifican" la población atómica en "familias", y entonces se dice que la función de línea es inhomogenea (sería el caso de un gas en el que el proceso dominante fuese un ensanchamiento por efecto Doppler).

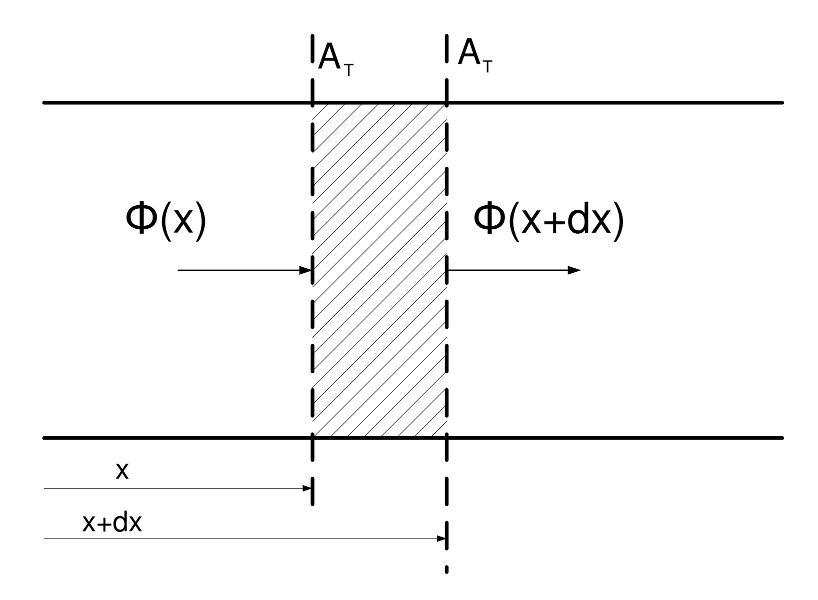


Figura 4.6: Volumen de control del medio amplificador.

y el proceso espontáneo (γ_{rad}), que es el siguiente (f es la frecuencia de la radiación y n el índice de refracción del medio):

$$\frac{B}{\gamma_{rad}} = \frac{c_{\circ}^{3}}{8\pi h f^{3} n^{3}} \tag{4.2}$$

También podemos determinar cómo varía con el tiempo la población atómica:

$$\left(-\frac{\partial N_2}{\partial t}\right)_{\omega} = \frac{N_2}{\tau_{esp}} l(\omega) d\omega + B(N_2 - N_1) W(\omega) l(\omega) d\omega \quad \Rightarrow$$

$$\left(-\frac{\partial N_2}{\partial t}\right)_{tot} = \frac{N_2}{\tau_{esp}} + B(N_2 - N_1) \int_0^{\infty} W(\omega) l(\omega) d\omega \tag{4.3}$$

4.3. Ganancia de un medio atómico

En este apartado pretendemos deducir la expresión del coeficiente $\alpha(\omega)$ teniendo en cuenta la emisión estimulada. Además, conociendo el valor de este coeficiente, podemos deducir las condiciones que deben verificarse para que el medio no atenúe la luz, sino que la amplifique. Evidentemente, lo que debemos hacer es determinar el número neto de fotones cedidos a la radiación que se propaga por el medio, es decir, la diferencia entre los añadidos por emisión estimulada menos los robados por absorción e imponer que este número sea positivo:

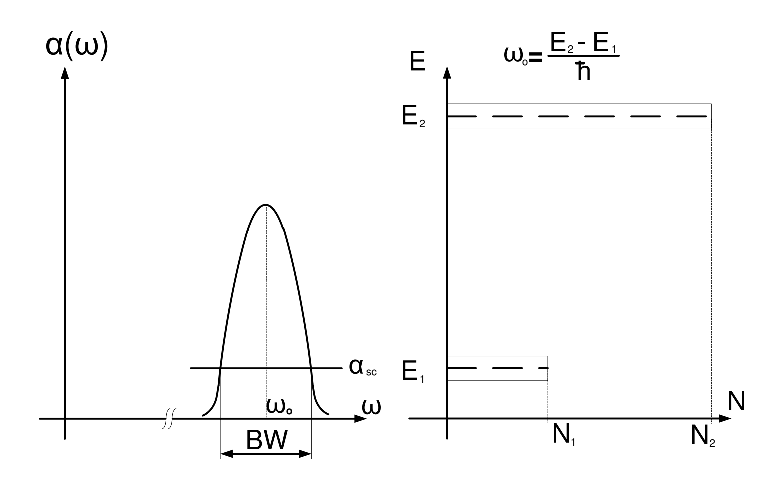


Figura 4.7: Condiciones de inversión de población en un medio atómico.

$$\left(-\frac{\partial N_2}{\partial t}\right)_{\omega} = \frac{\gamma_{rad} c_o^3}{8\pi h f^3 n^3}(N_2 - N_1) W(\omega) l(\omega) d\omega \tag{4.4}$$

De esta expresión deducimos que la condición para que el medio amplifique la radiación electromagnética es que el número de átomos que están en el nivel energético superior (el excitado) sea mayor que el número de átomos que están en el estado inferior (fundamental) [8]. Calculemos, a continuación, la evolución del flujo de fotones que atraviesan el medio atómico. En la fig. {4.6} se muestra un volumen de control que forma parte del medio por el que se propaga la luz. Este volumen de control está centrado en un punto genérico de la coordenada longitudinal x y está definido por dos secciones transversales de área A_t separadas entre sí por una distancia dx. El número total de fotones cedidos a la radiación en este volumen de control es la diferencia entre los añadidos por emisión estimulada menos los robados por absorción:

$$\left(-\frac{\partial N_2}{\partial t}\right)_{est-abs} \times A_t \times dx \tag{4.5}$$

Los fotones que se crean en este volumen de control hacen que por la superficie A_t centrada en $x + dx$ salga un número de fotones mayor que los que entran por la superficie A_t centrada en x. Es decir, podemos escribir:

$$\left(-\frac{\partial N_2}{\partial t}\right)_{est-abs} \times A_t \times dx = \Phi(x + dx) \times A_t - \Phi(x) \times A_t = -\frac{\partial \Phi}{\partial x} \times A_t \times dx \Rightarrow$$

$$\left(\frac{\partial N_2}{\partial t}\right)_{est-abs} = \frac{\partial \Phi}{\partial x} \tag{4.6}$$

Medio atómico con inversión de población

$N_2 - N_1 > 0$

Onda de pulsación cercana a ω_o propagándose

$E = E_o e^{[\alpha(\omega)-\alpha]z/2} e^{j[\omega t - \beta(\omega)z]}$

Figura 4.8: Onda propagándose por una medio atómico in inversión de población.

Pero ¿qué es el flujo exactamente? Pues el número de fotones que atraviesan la unidad de área transversal en la unidad de tiempo. El cociente $\frac{W(\omega)d\omega}{hf}$ es el número de fotones que, por unidad de volumen, hay en el intervalo espectral $(\omega, \omega + d\omega)$ y si lo multiplicamos por la velocidad con la que se mueve el haz $\frac{c_\circ}{n}$ obtenemos el número de fotones que atraviesan la unidad de área en la unidad de tiempo en el intervalo espectral $(\omega, \omega + d\omega)$, es decir, el flujo:

$$\Phi = \frac{W(\omega)c_\circ}{hfn} \tag{4.7}$$

La ganancia del medio por unidad de longitud (llamado coeficiente de atenuación si es un número negativo) es, por definición:

$$\alpha(\omega) := \frac{1}{\Phi}\frac{\partial \Phi}{\partial x} = \frac{\frac{\gamma_{rad} c_\circ^3}{8\pi h f^3 n^3}(N_2 - N_1)W(\omega)l(\omega)N_1 d\omega}{\frac{W(\omega)c_\circ}{hfn}} = \frac{\gamma_{rad}\lambda_\circ^2}{8\pi n^2}(N_2 - N_1)l(\omega) \tag{4.8}$$

Si ahora particularizamos para un medio que tenga como función de línea una lorentziana [9], el coeficiente de ganancia por unidad de longitud tendría la siguiente expresión, en la que el término que aparece en el numerador, $\frac{2}{\pi \Delta\omega}$, no es más que un coeficiente de normalización para conseguir que el área de la función de línea sea igual a la unidad y represente, por tanto, una densidad de probabilidad:

[8] En realidad la condición es que $g_2 N_2 - g_1 N_1 > 0$, donde g_1 y g_2 representan el grado de degeneración de los estados E_1 y E_2. Para simplificar, suponemos que $g_1 = g_2$.

[9] Es la función de línea típica de un medio atómico con ensanchamiento homogéneo.

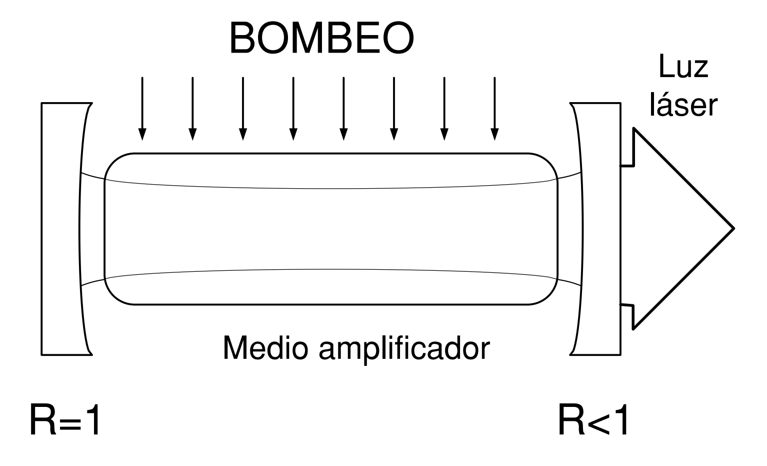


Figura 4.9: Cavidad resonante en un láser de niveles discretos.

$$\alpha(\omega) = \frac{\gamma_{rad}\lambda_\circ^2}{8\pi n^2}(N_2 - N_1)\frac{\frac{2}{\pi\Delta\omega}}{1 + (\frac{\omega - \omega_\circ}{\Delta\omega/2})^2} \tag{4.9}$$

Puesto que $\alpha(\omega)$ tiene unidades de la inversa de una longitud, es decir, $[\alpha(\omega)] = L^{-1}$, mientras que $N_2 - N_1$ tiene unidades de la inversa de un volumen, es decir, $[N_2 - N_1] = L^{-3}$, resulta claro por qué al término restante se le da el nombre de sección cruzada o sección eficaz. Su expresión analítica es la siguiente:

$$\sigma(\omega) = \frac{\gamma_{rad}\lambda^2}{8\pi}\frac{\frac{2}{\pi\Delta\omega}}{1 + (\frac{\omega - \omega_\circ}{\Delta\omega/2})^2} \tag{4.10}$$

4.4. Funcionamiento de un láser

Supongamos que tenemos un medio material en el que $N_2 - N_1 > 0$ (se dice entonces que en el medio hay una inversión de población). En este caso tenemos un coeficiente de ganancia por unidad de longitud $\alpha(\omega) > 0$. Al igual que en cualquier otro medio dieléctrico, existen fluctuaciones locales del índice de refracción que producen pérdida por scattering Rayleigh, pérdidas que podemos suponer constantes e independientes de la pulsación ω, tal como se muestra en la fig. {4.7}. Veamos qué es lo que ocurre cuando se propaga por este medio una radiación electromagnética de pulsación cercana a $\omega_\circ = \frac{E_2 - E_1}{\hbar}$.

En estas circunstancias, si la pulsación de la onda electromagnética que se propaga por el medio cae dentro del ancho de banda donde el coeficiente de amplificación es positivo, su amplitud crece a medida que se propaga, tal y como se muestra en la fig. {4.8}.

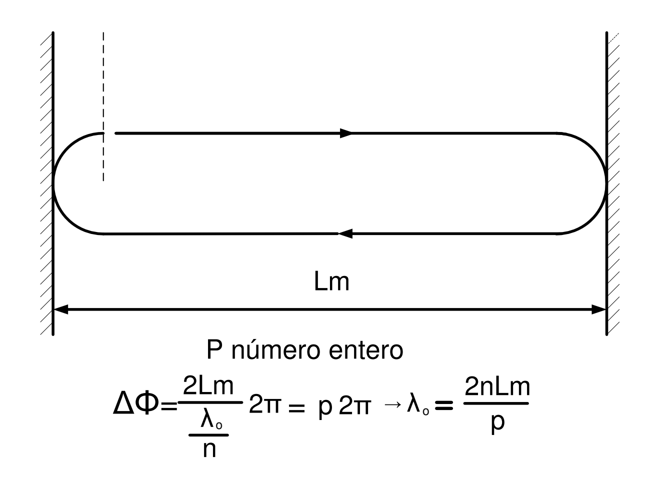


Figura 4.10: Longitudes de onda seleccionadas por la cavidad resonante.

Evidentemente, para mantener la inversión de población, esto es, $N_2 - N_1 > 0$, incluso en ausencia de radiación electromagnética, es preciso suministrar una determinada potencia al medio, ya que la tendencia natural de los átomos es caer del estado energético excitado (E_2) al fundamental (E_1) por emisión espontánea [10]. En ausencia de fenómenos estimulados y despreciando otros procesos de desexcitación, el tiempo que, en promedio, permanece un átomo en el estado superior antes de caer al estado fundamental es $\tau_{esp} = \frac{1}{\gamma_{rad}}$.

¿Cómo introducimos la realimentación en el láser? Limitamos el medio amplificador metiendo el gas atómico en un tubo cilíndrico hermético en cuyos extremos colocamos dos espejos (planos o curvos), paralelos entre sí. Uno de ellos debe tener un coeficiente de reflexión ligeramente inferior a la unidad, tal como se muestra en la fig. {4.9}. Mediante el mecanismo de bombeo (que aún no hemos descrito), se matiene la inversión de población. El objeto de la cavidad resonante es doble: por un lado, se consigue que un fotón [11] recorra el medio material varias veces antes de salir de la cavidad, de manera que compensamos el hecho de que el coeficiente de ganacia sea bajo [12]. Por otra parte la cavidad selecciona un conjunto discreto de longitudes de onda a las que el láser puede emitir. Estas longitudes de onda son aquellas para las que, después de un ida y vuelta completa por la cavidad, se produce un desfase un número entero de veces la cantida 2π. Esto se muestra en la fig. {4.10}.

La cavidad selecciona estas longitudes de onda porque sólo a éstas se produce una interferencia constructiva entre las ondas reflejadas por los espejos. Para el resto, la interferencia es destructiva por lo que desaparecen de la cavidad. Ahora lo que debemos plantearnos es por qué el láser acaba emitiendo a una única longitud de onda si existe todo un conjunto discreto de ellas para las que se cumple que $\alpha > \alpha_{sc} + \alpha_{esp}$, donde α_{esp} son las pérdidas que se producen por la reflexión parcial en el espejo con $R < 1$.

[10]En realidad hay procesos físicos por los que el átomo puede pasar del estado "2" al "1" sin emitir un fotón.

[11]Esta afirmación es poco rigurosa e incluso podríamos calificarla de errónea. Sin embargo, a veces resulta pedagógicamente útil referirse a los fotones como si fuesen bolas de billar.

[12]El medio activo está, generamente, muy diluido ya que la presión del gas que hay dentro de la cavidad es muy baja.

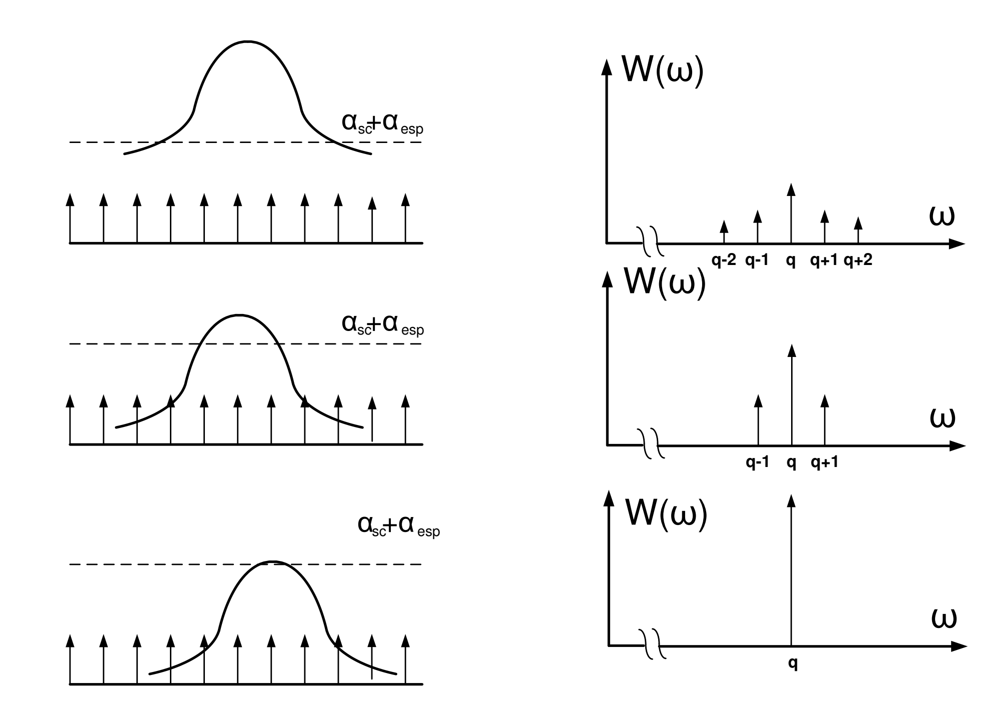


Figura 4.11: Evolución de la densidad volumétrica de energía electromagnética dentro de la cavidad. Las figuras de arriba representan un tiempo un poco posterior al arranque del láser y las de más abajo el régimen permanente.

En la fig. {4.11} vemos cómo evoluciona la densidad de energía dentro de la cavidad. Inicialmente hay generación de fotones por emisión espontánea. Una pequeña fracción de ésta coincide en frecuencia con los modos longitudinales de la cavidad para los que la ganancia por unidad de longitud es mayor que las pérdidas. De esta radiación hay, a su vez, otra pequeña parte se propaga perpendicularmente a los espejos de la cavidad. Esta pequeña fracción de emisión espontánea se realimenta positivamente y empieza a crecer su densidad volumétrica de energía. Esta situación se muestra en las figura de más arriba. A medida que crece $W(\omega)$, hay un mayor número de átomos que pasan del estado excitado al fundamental, con lo que la inversión de población disminuye y, consiguientemente, el coeficiente de ganancia también disminuye. Modos para los que inicialmente había realimentación positiva, pasan a ver un coeficiente de ganacia inferior al de pérdidas por lo que la densidad espectral de energía a esas frecuencias cae a cero. Esta situación se ha representado en las figura centrales de [4.11]. Por último se llega a una situación de régimen permanente, en la que pervive un único modo, para el que el coeficiente de ganacia compensa exactamente el de pérdidas. El medio amplificador repone los fotones que salen de la cavidad por los espejos o por scattering Rayleigh, pero no puede hacer crecer su número.

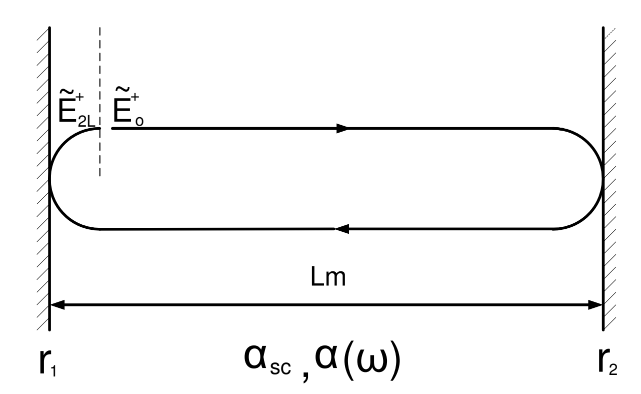


Figura 4.12: Condicion de oscilación del láser: después de una ida y vuelta por la cavidad resonante, la amplitud debe ser la misma y el desfase un número entero de veces la cantidad 2π.

Vemos que el láser tiene, en principio, tendencia e emitir a varias pulsaciones. Sin embargo hay una autorregulación a la baja de $\alpha(\omega)$, de manera que tras un cortísimo régimen transitorio emite a única pulsación [13] para la que se verifica que $\alpha(\omega)_{rp} = \sigma(\omega)(N_2 - N_1)_{rp} = \alpha_{sc} + \alpha_{esp}$.

4.5. Condiciones de oscilación de un láser

La condición de oscilación de un láser se obtiene imponiento condiciones estacionarias al campo electromagnético dentro de la cavidad, desechando, naturalmente, la condición trivial $E(z,t) = 0$. Para imponer condiciones estacionarias estudiamos la ida y la vuelta de una onda electromagnética e imponemos que no cambie ni la amplitud ni la fase de la onda como consecuencia de este recorrido. Ver fig. {4.12}.

$$E(z,t) = E_\circ exp[(\alpha(\omega) - \alpha_{sc})z]e^{-j\omega t} \quad t = \frac{zn}{c_\circ} \tag{4.11}$$

$$\begin{cases} E^+(0,0) = E_\circ \\ E^+(L, \frac{Ln}{c_\circ}) = E_\circ e^{[\alpha(\omega)-\alpha_{sc}]L} e^{-j\omega\frac{Ln}{c_\circ}} \\ E^-(L, \frac{Ln}{c_\circ}) = E_\circ r_2 e^{[\alpha(\omega)-\alpha_{sc}]L} e^{-j\omega\frac{Ln}{c_\circ}} \\ E^-(2L, \frac{2Ln}{c_\circ}) = E_\circ r_2 e^{[\alpha(\omega)-\alpha_{sc}]2L} e^{-j\omega\frac{2Ln}{c_\circ}} \\ E^+(2L, \frac{2Ln}{c_\circ}) = E_\circ r_1 r_2 e^{[\alpha(\omega)-\alpha_{sc}]2L} e^{-j\omega\frac{2Ln}{c_\circ}} \end{cases} \tag{4.12}$$

[13] En realidad esto es cierto sólo en medios en los que la función de línea es homogénea. Si no lo es, se puede producir el fenómeno de "hole burning" espectral [92] y el láser puede emitir a longitudes de onda correspondientes a varios modos longitudinales. De todas maneras, incluso si los mecanismos de ensanchamiento de la función de línea son homogéneos, puede producirse emisión láser a pulsaciones correspondientes a distintos modos longitudinales. Este fenómeno se conoce con el nombre de "hole burning" espacial y es especialmente importante en láseres a semiconductor [4].

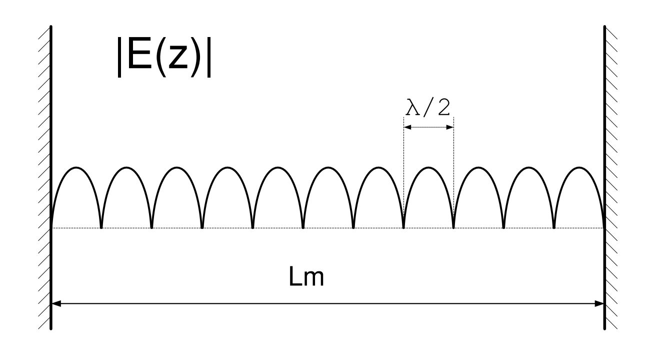


Figura 4.13: Estimación del orden del modo longitudinal presente en la cavidad de un láser de niveles discretos.

Para tener condiciones estacionarias debe verificarse:

$$\frac{E^+(2L,\frac{2Ln}{c_\circ})}{E^+(0,0)} = 1\angle p2\pi \quad p\in\mathcal{N} \Rightarrow \begin{cases} \alpha = \alpha_{sc} + \frac{1}{2L}Ln(\frac{1}{r_1}\frac{1}{r_2}) \\ \omega\frac{2Ln}{c_\circ} = p\times 2\pi \quad p\in\mathcal{N} \end{cases} \tag{4.13}$$

Estas dos últimas expresiones tienen una interpretación sencilla: En primer lugar, la ganacia del medio activo (por unidad de longitud) debe compensar las pérdidas (también por unidad de longitud) de la cavidad. De esta manera podemos tener una energía electromagnética constante dentro del cilindro y una potencia, también constante, saliendo por el espejo semireflectante. Por otro lado, como ya hemos visto, la cavidad selecciona un conjunto de pulsaciones o longitudes de onda determinadas: las correspondientes a los modos propios (o, también llamados, modos longitudinales [14]).

Vamos ahora a dar valores típicos a los parámetros de un laser de niveles discretos. En primer lugar, la cavidad resonante tiene una longitud del orden de $L \simeq 0{,}5m$, el índice de refracción es del orden de $n \simeq 1$ ya que el gas suele estar muy diluido (la presión es muy baja). El coeficiente de reflexión en potencia del espejo semireflectante es $R \simeq 0{,}8$ y un coeficiente de scattering Rayleigh típico puede ser $\alpha_{sc} \simeq 1cm^{-1}$. Teniendo en cuenta que $R = |r|^2$ resulta que la ganacia del medio debe ser del orden de $\alpha \simeq 4cm^{-1}$. Sabemos también que las pulsaciones propias de la cavidad vienen dadas por la expresión $\omega_p = p\times 2\pi\frac{c_\circ}{2\pi L}$, de manera que la separación entre dos modos consecutivos es del orden de $\Delta\omega = \omega_{p+1}-\omega_p = 2\pi\frac{c_\circ}{2nL} \simeq 2\pi\times 300\times 10^6$. Viendo la fig. {4.13} podemos estimar cuál es el modo longitudinal existente en la cavidad resonante. Para ello, tenemos que tener en cuenta que el cociente $\frac{L_m}{\lambda/2} = p$ es un número entero. Dado que la longitud es del orden de 50cm, la longitud de onda de emisión es del orde de $1\mu m$, resulta que $p \simeq 5\times 10^5$.

[14] Estamos suponiendo que la distribución transversal admisible para campo electromagnética en la cavidad es única. Normalmente los láseres de niveles discretos y semiconductores (salvo los que se optimizan para emitir la máxima potencia posible) se diseñan para que esto sea así.

4.6. Mecanismo de bombeo.

Existe una gran variedad de medios activos con los que se puede fabricar un láser. Sin embargo todos, salvo el láser a semiconductor, tienen una característica común: el medio activo está muy diluido, es decir, hay pocos átomos ópticamente activos (también pueden ser moléculas o iones) por unidad de volumen. Si tenemos en cuenta que $\alpha(\omega)$ es proporcional a $l(\omega)$ (función de densidad de probabilidad de transición entre estados) resulta evidente que la amplificación será tanto más sencilla cuanto menor sea la anchura espectral de la ganancia [15]. Esto sólo es posible en medios diluidos. En los láseres de estado sólido (salvo los láseres semiconductores), los elementos activos son iones dispersos en el seno de un medio vítreo que actúa de soporte, en los láseres líquidos (de colorantes) ocurre lo mismo: el líquido no es más que un medio donde se disuelve un colorante, que es el material ópticamente activo. Por último, en los gases, la presión de trabajo es muy baja, de forma que los átomos o moléculas están separados por distancias medias grandes.

4.6.1. Láser de tres niveles.

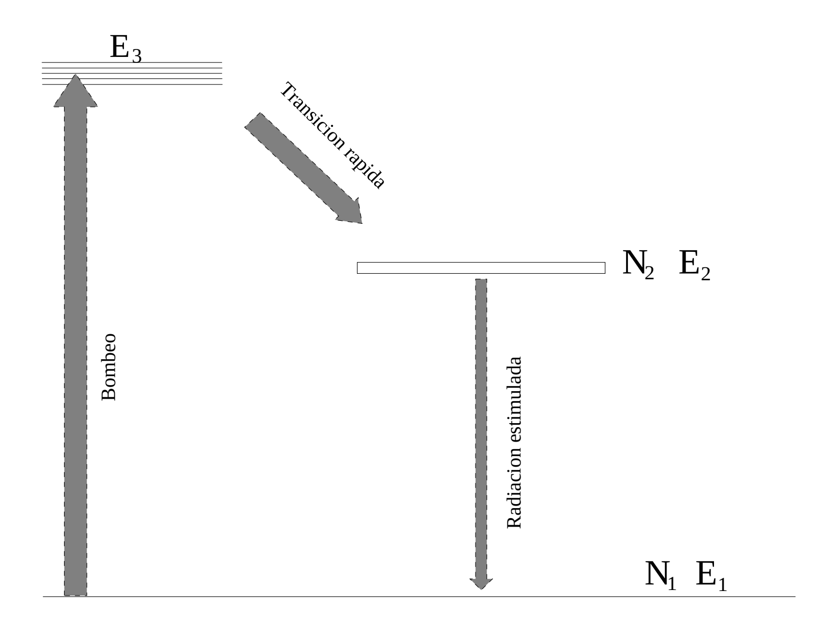


Figura 4.14: Medio activo de tres niveles.

Las características de un láser de tres niveles se muestran en la fig. {4.14}. El nivel energético "1" es el estado fundamental de los átomos, mientras que los estados "2" y "3" son estados excitados. El sistema de bombeo se encarga de ceder energía al medio amplificador, induciendo transiciones entre el estado fundamental y el "3", siendo éste inestable, es decir, su tiempo medio de permanencia es muy pequeño, es decir, $\tau_{32} \ll \tau_{21}$. Por este motivo los átomos se desexcitan rápidamente al estado "2", donde permanecen más tiempo al ser un estado metaestable.

[15]Para que exista emplificación es necesario que $\alpha(\omega) > \alpha_p$. Si la anchura espectral de $l(\omega)$ es grande, su amplitud debe ser pequeña (de hecho resulta habitual aproximar $l(\omega) \simeq \frac{1}{\Delta\omega}$) y la amplificación difícil.

Entre los niveles ”1” y ”2” se producen los fenómenos de absorción y emisión estimulada, que son los responsables de la amplificación de luz. En este tipo de láseres el bombeo tiene que ser muy enérgico, puesto que hay que llevar un número importante de átomos al estado ”2” (al menos la mitad). Teniendo en cuenta que el bombeo es generalmente óptico (en este tipo de sistemas) los láseres a tres niveles sólo son factibles si en la transición ”1”–”3” hay una banda de absorción muy importante [16].

4.6.2. Láseres de cuatro niveles.

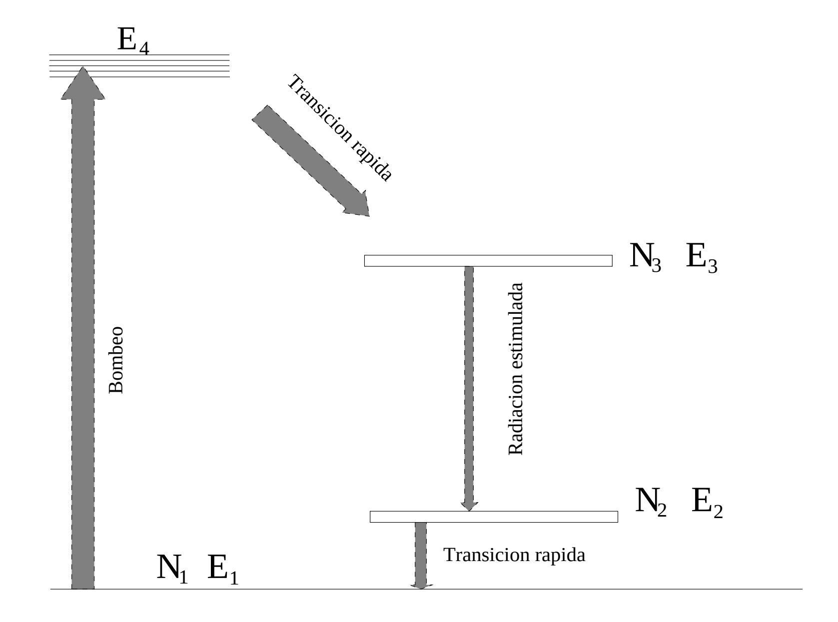


Figura 4.15: Medio activo de cuatro niveles.

En la fig. {4.15} se muestra la estructura energética de un láser a cuatro niveles. El nivel 1 es el fundamental y el bombeo se produce entre los niveles ”1” y ”4”. El nivel ”4” es inestable, por lo que rápidamente se desexcita hacia el nivel ”3”, donde el tiempo medio de permanencia es, en términos relativos, grande. Entre los niveles ”3” y ”2” se producen los fenómenos de emisión estimulada y de absorción. Los átomos que pasan al nivel ”2” (por emisión estimulada) se desexcitan rápidamente, ya que se trata de un estado inestable, pasando rápidamente al nivel fundamental. Por consiguiente la población atómica del nivel ”2” es muy pequeña. Vemos que con un pequeña potencia de bombeo se consigue la inversión de población, ya que $N_2 \simeq 0$ y se puede producir amplificación óptica. Esta es la gran ventaja de los sistemas de cuatro niveles y el motivo por el que la mayor parte de los láseres de niveles discretos son de este tipo.

[16]En realidad el estado ”3” no es un único nivel energético, sino un conjunto de niveles próximos entre sí que posibilitan la absorción de una parte significativa de los fotones que crea la fuente de luz encargada del bombeo (normalmente una lámpara flash o una láser semiconductor).

4.7. Problemas

Problema 6: Se desea modificar la ecuación de polarización macroscópica de un gas de manera que tenga en cuenta la emisión estimulada. Sugerencia: compare los . coeficientes de ganacia (en realidad de atenuación) que se derivan del modelo clásico de osciladores y fuerce a que coincida con la expresión obtenida en este capítulo.

Solución: La ecuación de la polarización macroscópica deducida en el segundo tema es:

$$\ddot{P}_x + (\gamma + \frac{2}{T_o})\dot{P}_x + \omega_o^2 P_x = \frac{Nq^2 E(t)}{dVm}$$

de la que se deducía el siguiente coeficiente de absorción:

$$\alpha(\omega) = \frac{2\pi}{\lambda_\circ}\mathcal{I}\{\chi(\omega)\} = \frac{\pi}{\lambda_\circ}\frac{q^2}{m\epsilon_\circ\omega_\circ}\frac{N}{dV}\frac{\frac{2}{\Delta\omega}}{(\frac{\omega_\circ - \omega}{1 + \Delta\omega/2})^2}$$

la expresión que hemos obtenido en este tema es:

$$\alpha(\omega) = \frac{\gamma_{rad}\lambda_\circ^2}{8\pi n^2}(N_2 - N_1)\frac{\frac{2}{\pi\Delta\omega}}{1 + (\frac{\omega - \omega_\circ}{\Delta\omega/2})^2}$$

Vemos, pues, que lo que hay que hacer es identificar la densidad de osciladores clásicos $\frac{N}{dV}$ con la densidad de átomos en inversión de pobleción $N_2 - N_1$ y el cociente $\frac{q}{m}$ con $\frac{\gamma_{rad}\lambda_\circ^3\epsilon_\circ\omega_\circ}{8\pi^2 n^2}$.

Con estas sustituciones queda:

$$\ddot{P}_x + (\gamma + \frac{2}{T_\circ})\dot{P}_x + \omega_\circ^2 P_x = \frac{\gamma_{rad}\lambda_\circ^3\epsilon_\circ\omega_\circ}{8\pi^2 n^2}(N_2 - N_1)E(t)$$

Que es la expresión mecanocuántica correcta de la polarización macroscópica de un medio dieléctrico de dos niveles.

Capítulo 5

Ganacia óptica de un semiconductor

Resumen: En este tema vamos a estudiar el medio activo del transmisor que más se emplean en los sistemas de comunicación por fibra óptica: el láser semiconductor.
Los semiconductores son sólidos monocristalinos cuyas propiedades eléctricas se sitúan en un punto intermedio entre las de los metales y los aislantes. En este capítulo se dará una breve introducción a las propiedades físicas de los semiconductores para estudiar, seguidamente, los procesos de interacción luz–materia en estos materiales. En particular pretendemos calcular el coeficiente de ganancia (o de atenuación) y relacionarlo con la concentración de portadores. La materia tratada en este capítulo es fundamental para entender el funcionamiento de los láseres de cavidad Fabry-Pérot y los de cavidad distribuida.

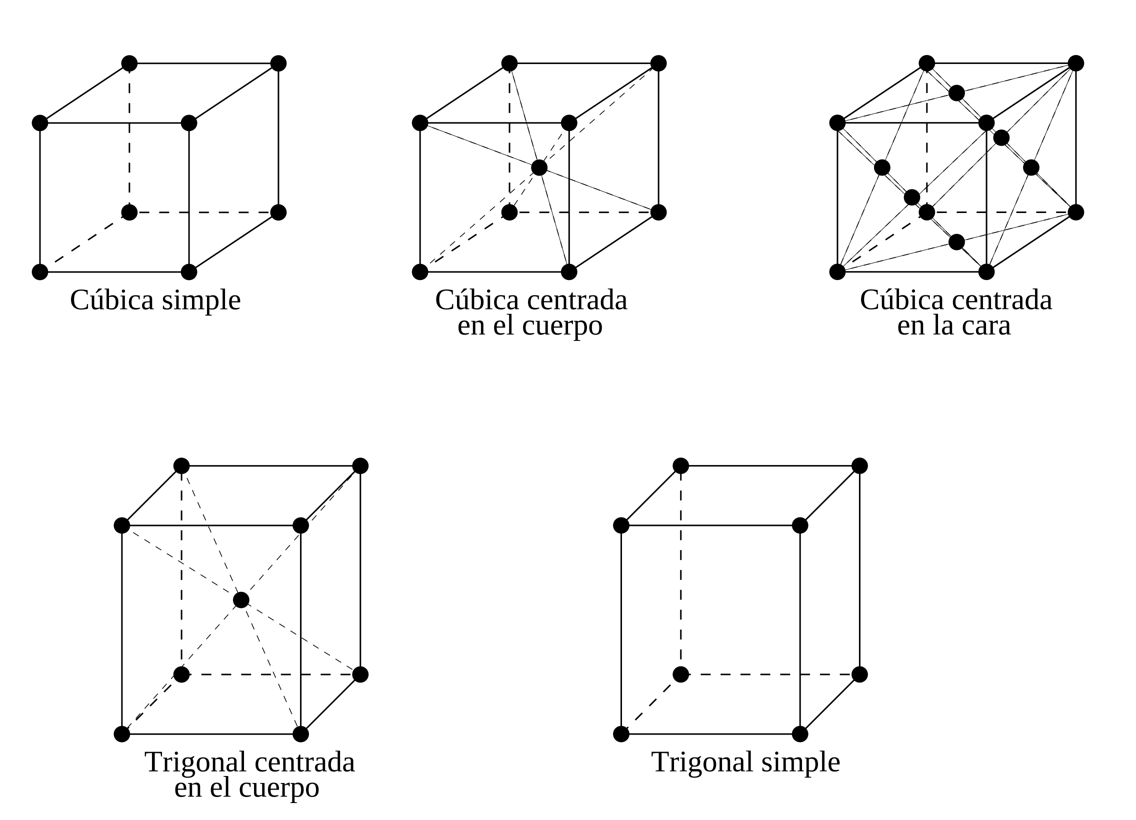


Figura 5.1: Algunos ejemplos de celdillas elementales.

5.1. Introducción

Los sólidos cristalinos [12]–[88] tienen un estructura muy regular. Los átomos que los constituyen se ordenan en el espacio de una manera periódica casi perfecta, de forma que el sólido puede verse como un "mosaico", en el que existe un bloque elemental que se repite regularmente en las tres direcciones del espacio [93]. En la fig. {5.1} se muestran distintos ejemplos de celdas elementales, a partir de las cuales puede construirse el sólido cristalino desplazando ésta a lo largo los tres vectores de translación asociados a sus aristas. Estos cristales son sustancias *anisótropas*, es decir, sustancias cuyas propiedades dependen de la dirección o el plano en que se midan. No es difícil convencerse de esto si se observa que existen direcciones y planos esencialmente distintos, donde la periodicidad en la repetición de átomos es distinta.

La extensión espacial de las "órbitas" [1] de los electrones de un átomo aislado es tanto mayor cuanto mayor es la energía del electrón. Por ello, cuando se aproximan entre sí los átomos para formar un sólido cristalino, son precisamente los electrones de valencia (que orbitan a mayor distancia del núcleo) los únicos que interaccionan [2]. Esto justifica el considerar separadamente el conjunto formado por el núcleo del átomo y los electrones internos (a este conjunto lo llamaremos núcleo de la red cristalina) por un lado y los electrones de valencia por otro.

Cuando dos átomos iguales están lo suficientemente separados, los electrones de uno

[1]En realidad se debería hablar de las funciones de onda asociadas [14] a los electrones en estados estacionarios (es decir, de energía constante).
[2]En algunos elementos esta hipótesis simplificativa no es válida.

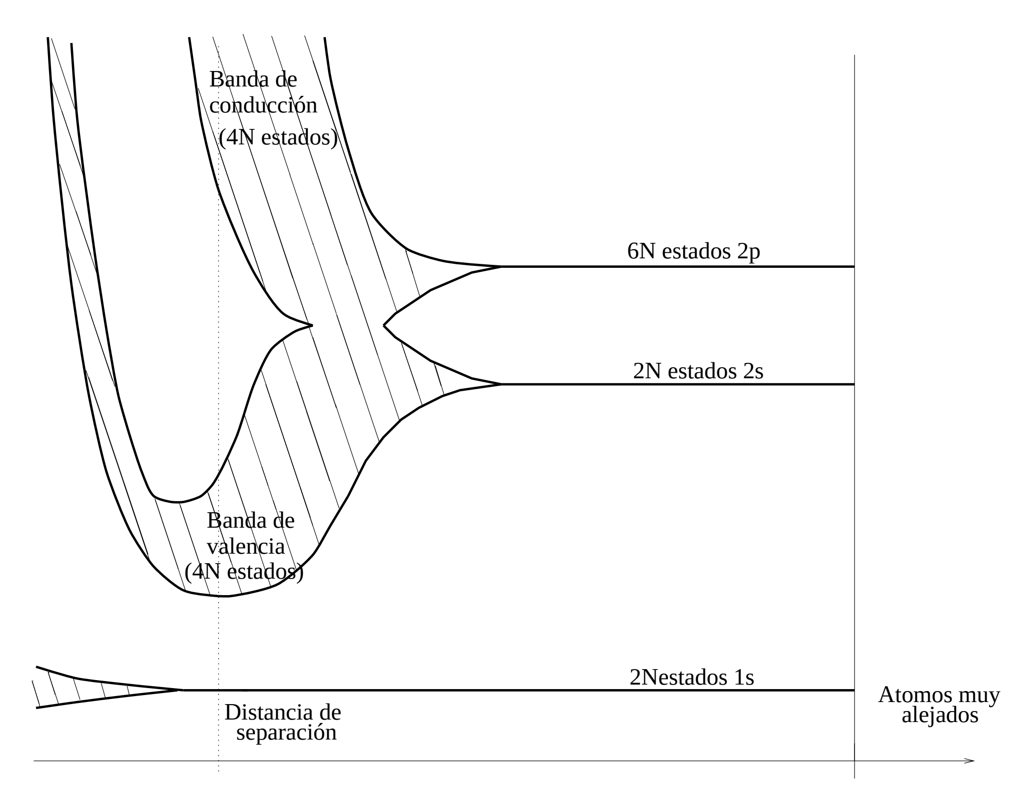


Figura 5.2: Formación de las bandas de valencia y conducción de un cristal de carbono.

y otro no interaccionan entre sí. Las "órbitas" correspondientes a electrones equivalentes de estos dos átomos no solapan. Aquellos electrones que ocupan niveles similares en ambos átomos tienen la misma energía y sus "órbitales" son iguales (mientras no haya interacción entre ambos). Sin embargo, a medida que se aproximan, las "órbitas" comienzan a solapar con lo que se ven perturbadas. Cualitativamente se observa un desplazamiento de los niveles energéticos de ambos electrones de forma que éstos no coincidan.

En la fig. {5.2} se observa lo que ocurre a medida que se acercan los N átomos de una red cristalina de carbono [71]. Los 2N niveles energéticos 2s de los átomos, así como los 6N niveles 2p, empiezan a separarse, formando, inicialmente, dos "bandas" de niveles energéticos, una correspondiente a los 2N niveles 2s y otra correspondiente a los 6N niveles 2p (N es el número de átomos de la red cristalina). Si el acercamiento entre átomos continúa, las dos bandas se funden en una sola. Si la aproximación es aún mayor, esta banda se divide en dos, cada una de las cuales contiene 4N niveles energéticos. La banda inferior (formada por orbitales "enlazantes") recibe el nombre de *banda de valencia* y la superior (formada por los orbitales "antienlazantes") *banda de conducción*. En la figura se ha representado también la distancia de separación entre átomos en la red del diamante, es decir, la distancia para la cual la energía total del sistema se hace mínima. El intervalo de separación entre ambas bandas recibe el nombre de *banda prohibida* y el valor de la energía que las separa, *energía de la banda prohibida*.

A una temperatura igual a 0K, los electrones ocupan los niveles de energía más bajos. El diamante cuenta con cuatro electrones en su última capa, por lo que la banda de valencia estará totalmente ocupada, mientras que la banda de conducción estará totalmente vacía. Para que un electrón pueda desplazarse como una partícula libre a lo largo de la red cristalina

(condición indispensable para que sea conductor de la electricidad) debe tener estados de energía superiores accesibles (vacíos) dentro de la banda a la que pertenece [3]. Por ello se puede deducir que a 0K el diamante es un *aislante perfecto*. A temperaturas superiores a 0K los electrones no tienen por qué ocupar necesariamente los niveles energéticos más bajos posibles y existe una probabilidad no nula de que un electrón que pertenecía a la banda de valencia ocupe un estado energético de la banda de conducción. Esta probabilidad depende tanto de la temperatura como de la anchura de la banda prohibida. En el diamante la anchura de la banda prohibida es lo suficientemente grande como para que a cualquier temperatura razonable la probabilidad de que esto ocurra sea ínfima. Por ello el diamante es una excelente aislante incluso a temperaturas muy altas.

Los metales son buenos conductores de la electricidad debido a que la banda de conducción no está vacía sino que se encuentra parcialmente llena, o bien puede ocurrir que la banda de valencia y la banda de conducción solapen entre sí (este es el caso de los metales de transición).

Hay determinados sólidos cristalinos que presentan características de conducción eléctrica intermedias entre las de los aislantes y los metales. Son los llamados *semiconductores*, entre los que cabe destacar el Silicio, el Germanio y el Arseniuro de Galio y sus derivados [1]. En estos materiales la anchura de la banda prohibida es lo suficientemente estrecha como para que a temperaturas razonables haya una concentración no despreciable de electrones en la banda de conducción. Este tipo de materiales son buenos aislantes a temperaturas bajas y su conductividad eléctrica crece a medida que aumenta la temperatura.

5.2. Portadores libres de carga en los semiconductores.

A medida que la temperatura de un semiconductor aumenta por encima de los 0K, algunos electrones de valencia adquieren la suficiente energía térmica como para pasar a la banda de conducción. Cuanto mayor sea la temperatura, mayor será la probabilidad de que un electrón adquiera una energía térmica igual o superior a la anchura de la banda prohibida, con lo que el número de electrones en la banda de conducción aumentará. Estos electrones, al saltar a la banda de conducción, dejan estados energéticos libres en la banda de valencia, la cual estaba completamente llena a 0K. Los estados libres de la banda de valencia reciben el nombre de *huecos*. Tanto los electrones como los huecos son portadores de carga (los primeros de carga negativa y los segundos de carga positiva) que se desplazan "libremente" [4] por las bandas de conducción y valencia respectivamente, con lo que contribuyen al fenómeno de conducción de la electricidad [71]–[93].

En esta sección se pretende determinar la concentración de portadores de carga (electrones y huecos) que, a una temperatura dada, existe en un semiconductor. Primeramente

[3]Los electrones de una banda totalmente ocupada, si se encuentran sometidos a un campo eléctrico, realizan un movimiento armónico, llamado oscilación de Bloch, alrededor de los núcleos de la red cristalina.

[4]La hipótesis de partícula libre es aproximadamente cierta sólo en las cercanías del mínimo de la banda de conducción y del máximo de la banda de valencia en semiconductores de transición directa. En los de transición indirecta la hipótesis deja de ser válida en la banda de conducción. Sin embargo, gracias a la presencia de varios mínimos simétricos en la primera zona de Brillouin [12], al promediar el comportamiento de los electrones, su movimiento se puede describir como si fuesen libres.

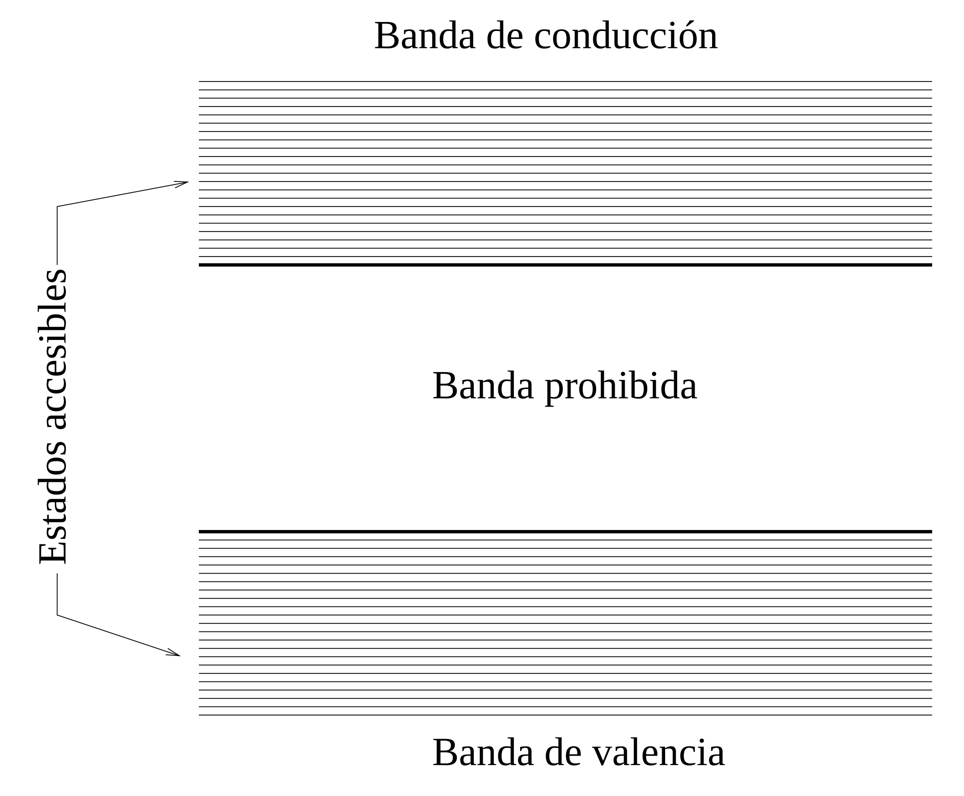


Figura 5.3: Discretización de estados energéticos dentro de las bandas de valencia y conducción.

se centrará el estudio sobre los semiconductores *intrínsecos*, es decir, aquellos en los que electrones y huecos se crean exclusivamente mediante excitación térmica. Estos portadores de carga reciben el nombre de *portadores intrínsecos* y la conductividad debida a estos portadores (la única que existe en los semiconductores intrínsecos), conductividad intrínseca. En un semiconductor intrínseco el número total de electrones en la banda de conducción es igual al número total de huecos en la banda de valencia ya que, al pasar un electrón de la banda de valencia a la de conducción, automáticamente éste deja un hueco en la banda de valencia.

Para hallar la conductividad de un semiconductor hay que determinar previamente dos funciones: La *densidad de estados* energéticos accesibles en las bandas de conducción y valencia [12] y la *probabilidad de ocupación* [12], a una temperatura dada, de estos niveles energéticos.

A partir del análisis hecho para los semiconductores intrínsecos, se estudiarán los semiconductores extrínsecos, es decir, aquellos cuyas propiedades eléctricas dependen de impurezas (llamadas dopantes) que se añaden al semiconductor.

5.2.1. Densidad de estados en las bandas de valencia y conducción.

Los niveles de energía permitidos en un sólido cristalino son discretos, aunque están separados entre sí por saltos de energía extremadamente pequeños, tal y como se muestra en la fig. {5.3}. Por esta razón, en general, se supone que forman un continuo de valores. Consideremos un nivel de energía situado en la banda prohibida. Según se dijo en apartados

anteriores, un electrón de un semiconductor ideal no puede tener ese valor de energía y, por tanto, si tomamos un intervalo estrecho de energías alrededor de ese valor, resulta que el número de estados energéticos (por unidad de volumen) accesibles en dicho intervalo es nulo y, por consiguiente, la densidad de estados, es decir el número de estados accesibles dividido por la anchura del intervalo considerado y por el volumen del semiconductor, es, en la banda prohibida, igual a cero. Si consideramos una energía superior al mínimo de la banda de conducción, resulta que ese estado sí es accesible para un electrón y, por consiguiente, la densidad es mayor que cero. La separación entre estos niveles no es constante, sino que es menor para los niveles energéticos más altos, es decir, la densidad de niveles energéticos es más alta (en la banda de valencia ocurre justo lo contrario: cuanto menor es la energía mayor es la densidad de estados). Las expresiones de las densidades de estados de los electrones en las cercanías del mínimo de la banda de conducción y de los huecos en las cercanías del máximo de la banda de valencia son, aproximadamente [101]:

$$
\begin{aligned}
N_e(E) &= \frac{8\sqrt{2}\pi}{h^3} m_e^{\frac{3}{2}} \sqrt{E - E_c} \\
N_h(E) &= \frac{8\sqrt{2}\pi}{h^3} m_h^{\frac{3}{2}} \sqrt{E_v - E}
\end{aligned}
\tag{5.1}
$$

donde E_c es la energía del mínimo de la banda de conducción, E_v es la energía del máximo de la banda de valencia y m_e y m_h son las masas efectivas de electrones y huecos en las bandas de conducción y valencia. La expresiones (5.1) ponen de manifiesto que, a medida que nos adentramos en la banda de conducción, el número de estados accesibles para los electrones crece con la raíz cuadrada de la energía. La afirmación análoga es cierta para los huecos en la banda de valencia. El coeficiente de proporcionalidad no es en ambos casos el mismo ya que, en general, las masas efectivas de electrones y huecos no son iguales.

5.2.2. Probabilidad de ocupación de estados

Para determinar el número de portadores de carga en la banda de conducción y en la banda de valencia necesitamos conocer, además de la densidad de estados energéticos accesibles $N(E)$, la *probabilidad* de que, supuesto que sea accesible, el estado de energía E esté ocupado. Esta característica viene definida por una función $f(E,T)$ de distribución donde E varía de 0 a infinito. El sentido de esta función es claro. Supongamos un estado de energía E; supuesto que este valor de energía pertenece la la banda de conducción (estado accesible de energía) la probabilidad de que esté ocupado por un electrón es $f(E,T)$.

La expresión de $f(E,T)$ es la siguiente:

$$
f(E,T) = \frac{1}{1 + e^{\frac{(E-E_f)}{kT}}}
\tag{5.2}
$$

Esta función recibe el nombre de función de distribución de Fermi–Dirac [101] y describe el comportamiento estadístico de cualquier población de fermiones y en particular el de los electrones y huecos en un sólido critalino. E_f es una "constante" en el sentido de que no es función de E, pero depende, aunque sea débilmente, de la temperatura y de las

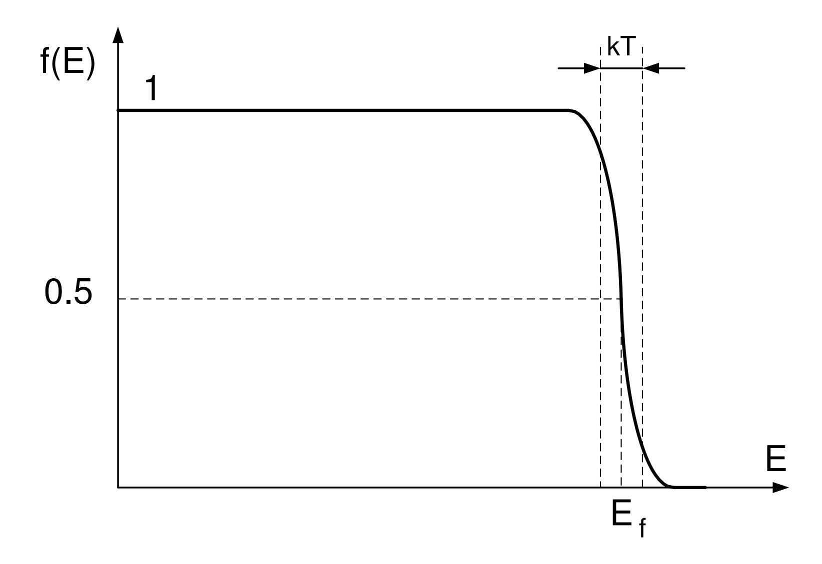


Figura 5.4: Función de distribución de Fermi-Dirac.

masas efectivas de los portadores. E_f recibe los nombres de potencial químico, nivel de Fermi o energía de Fermi.

En la fig. {5.4} hemos representado la función de distribución de Fermi–Dirac para una temperatura dada. En ella se observa que un estado de energía igual a la del nivel de Fermi tiene una probabilidad 0.5 de estar ocupado. En una sección posterior veremos cómo se determina E_f y cómo depende de la temperatura. La constante k es la constante de Boltzmann.

5.2.3. Concentración de electrones y huecos en equilibrio térmico en un semiconductor intrínseco.

La concentración de electrones en la banda de conducción, es decir, el número de electrones que hay por unidad de volumen en un intervalo infinitesimal de energía $(E, E+dE)$, se obtiene multiplicando la función de densidad de estados en la banda de conducción por la probabilidad de que dichos estados estén ocupados, es decir, por la función de distribución de Fermi-Dirac $f(E,T)$. Integrando este producto para energías superiores al mínimo de la banda de conducción se obtiene la concentración total de electrones libres en la banda de conducción [77].

La concentración de huecos en la banda de valencia se obtiene multiplicando la función de densidad de estados en la banda de valencia por la probabilidad de que dichos estado no estén ocupados (ésta es, precisamente, la probabilidad de que exista un hueco), es decir por $1 - f(E,T)$. Integrando este producto para energías inferiores al máximo de la banda de valencia se obtiene la concentración total de huecos libres en la banda de valencia.

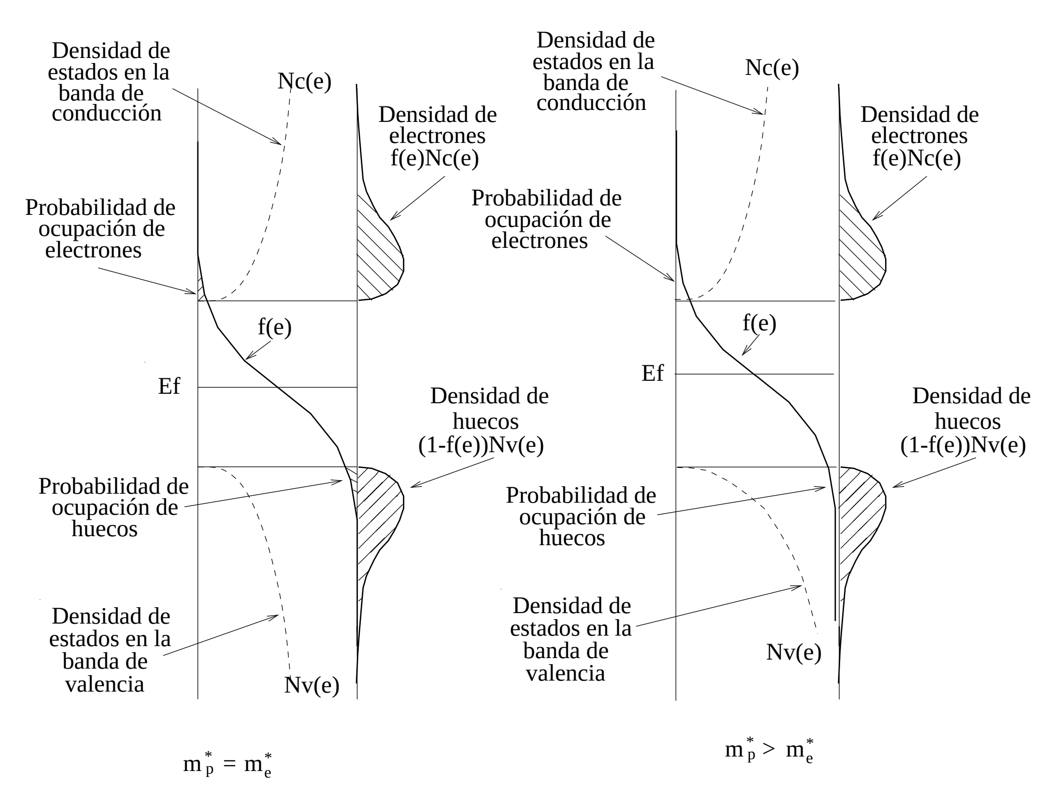


Figura 5.5: Concentración de electrones y huecos en un semiconductor.

En condiciones de equilibrio térmico, el número de electrones dn que, por unidad de volumen, tienen una energía dE alrededor de E en la banda de conducción es:

$$dn_o = N_e(E)f(E,T)dE = \frac{8\sqrt{2}\pi}{h^3} m_e^{\frac{3}{2}} \sqrt{E - E_c} \frac{1}{1 + e^{\frac{(E-E_f)}{kT}}} dE \tag{5.3}$$

Esta es la expresión analítica del diferencial de una de las curvas rayadas que se muestran en la fig. {5.5} (la correspondiente a la banda de conducción). Para obtener la concentración total de electrones en la banda de conducción habría que integrar para valores de E que van de E_c a ∞. Sin embargo, la función (5.3) no tiene una primitiva, ni puede calcularse de forma exacta (por otros medios como, por ejemplo, el método de los residuos) esta integral definida, por lo que tendríamos que realizar una integración numérica. Sin embargo, si la energía de Fermi está dentro de la banda prohibida a varias unidades kT lejos del borde de la banda de conducción (es decir, si $E_c - E_f \gg kT$), entonces podemos aproximar la función de distribución de Fermi-Dirac en la banda de conducción mediante [71]:

$$f(E,T) \simeq e^{-\frac{(E-E_f)}{kT}} \tag{5.4}$$

Las condiciones para que esta aproximación sea válida se verifican en todos los semiconductores intrínsecos normalmente utilizados en la fabricación de dispositivos. Esta aproximación recibe el nombre de aproximación de Boltzmann y tiene una interpretación física sencilla: la estadística de electrones y huecos pasa de ser una estadística cuántica a ser

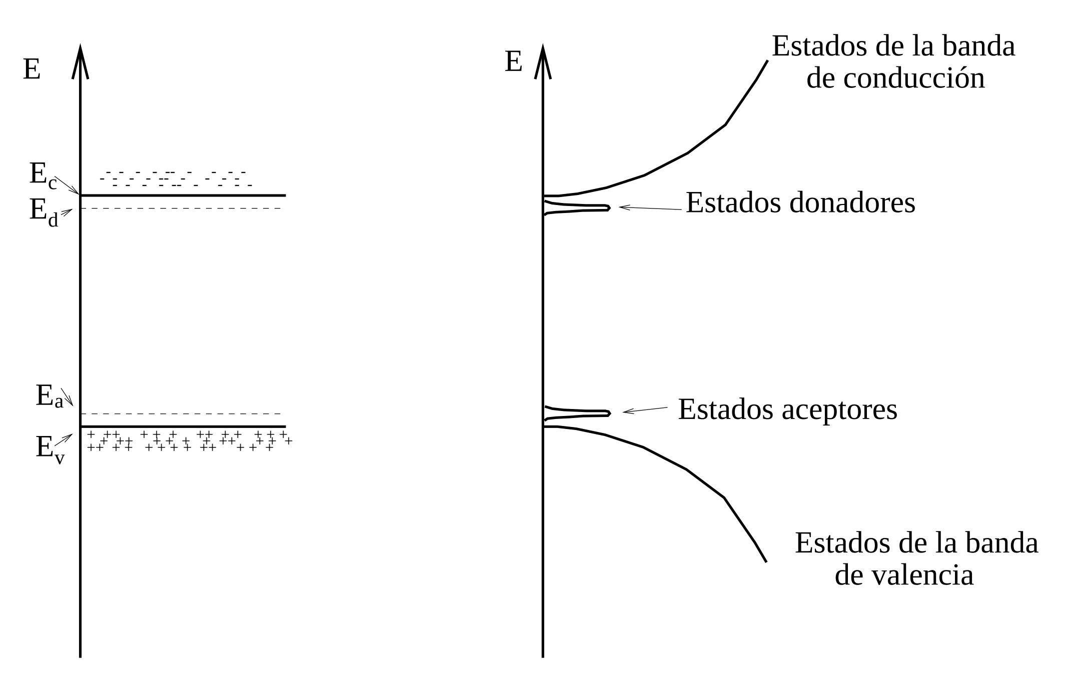


Figura 5.6: Estados energéticos donadores y aceptores. En un semiconductor dopado se da una de las dos posibilidades: o bien existen estados donadores, o bien existen estados aceptores.

una estadística clásica, ya que el número de portadores es pequeño y el principio de exclusión de Pauli no juega un papel importante.

Efectuando la integral antes mencionada, resulta:

$$n_i \simeq U_c e^{-\frac{(E_c - E_f)}{kT}}$$
$$U_c = 2\left(\frac{2\pi m_e kT}{h^2}\right)^{\frac{3}{2}} \tag{5.5}$$

El número de huecos p_i por unidad de volumen en la banda de valencia de un semiconductor, puede calcularse de un modo similar y se obtiene:

$$p_i \simeq U_v e^{-\frac{(E_f - E_v)}{kT}}$$
$$U_v = 2\left(\frac{2\pi m_h kT}{h^2}\right)^{\frac{3}{2}} \tag{5.6}$$

Si m_h y m_e fuesen exactamente iguales (esto no ocurren en ningún semiconductor de tipo IV o III-V [1]), las curvas de densidad de estados en las dos bandas son totalmente simétricas, por lo que la energía de Fermi debe quedar exactamente en el centro de la banda prohibida. De otra forma, la población de electrones en la banda de conducción y la de huecos en la banda de valencia no coincidirían. Si m_h y m_e no son iguales (normalmente la masa efectiva de los electrones es más ligera que la de los huecos), el nivel de Fermi debe desplazarse ligeramente hacia arriba o hacia abajo de forma que al efectuar las integrales se

obtengan valores iguales. En este caso, el nivel de Fermi queda cerca del centro de la banda prohibida pero no exactamente en él. Esto se muestra en la fig. {5.5}.

En un semiconductor intrínseco las concentraciones de electrones y huecos son iguales, por lo que identificando (5.5) con (5.6) puedemos deducir la dependencia del nivel de Fermi con la temperatura. Aconsejamos al lector que demuestre la siguiente expresión, cierta para un semiconductor intrínseco:

$$E_f = \frac{1}{2}(E_v + E_c) + \frac{3}{4}kT \ln \frac{m_h}{m_e} \tag{5.7}$$

Dado que, a temperaturas ordinarias, kT es una energía muy pequeña en comparación con la anchura de la banda prohibida (del orden de magnitud de $1eV$), resulta que, aún en el caso de semiconductores en los que las masas efectivas de los portadores de carga no sean iguales, el nivel de Fermi está situado aproximadamente en el centro de la banda prohibida (en un semiconductor intrínseco). Si $m_e < m_h$ el nivel de Fermi se desplaza ligeramente hacia el mínimo de la banda de conducción y si $m_h < m_e$, se desplaza ligeramente hacia el máximo de la banda de valencia.

Se puede demostrar con facilidad [2] que el producto $n_o p_o$ es una función exclusiva de la anchura de la banda prohibida, las masas efectivas y la temperatura, siendo independiente del nivel de Fermi. Los semiconductores extrínsecos pueden describirse (tal como veremos en la próxima sección) considerando que su nivel de Fermi está desplazado en relación a la posición que tendría si fuese intrínseco. Al no depender $n_\circ p_\circ$ del nivel de Fermi, la expresión que sigue, calculada tomando $n_o = n_i$ y $p_o = p_i$, es válida también para los semiconductores extrínsecos. (Con el subíndice $\circ$ se denota una situación de equilibrio térmico).

$$n_\circ p_\circ = n_i^2(T) = p_i^2(T) = 4\left(\frac{2\pi kT}{h^2}\right)^3 (m_e m_h)^{\frac{3}{2}} e^{-\frac{E_g}{kT}} \tag{5.8}$$

La demostración de esta fórmula la dejamos como ejercicio para el lector. Este resultado, que pone de manifiesto que el producto de la concentración total de electrones y huecos es independiente del dopado de un semiconductor extrínseco, se va a usar con mucha frecuencia.

5.2.4. Concentración de electrones y huecos en equilibrio térmico en un semiconductor extrínseco.

Es muy fácil introducir cantidades muy pequeñas de substancias tales como el arsénico [5], u otros elementos pertenecientes al grupo V de la tabla periódica, en cristales puros de silicio o germanio, como impurezas de substitución, es decir, como átomos de impurezas que ocupan sitios de la red cristalina que normalmente estarían ocupados por átomos del semiconductor de tipo IV. Los átomos del grupo V tienen cinco electrones de valencia.

[5]Estamos suponiendo que el semiconductor de tipo IV (cuarta columna del sistesma periódico) como, por ejemplo, el silicio, aunque todos los dispositivos ópticos se fabrican con semiconductores de tipo III–IV, es decir, compuestos de elementos de las columnas tercera y quinta del sistema periódico como, por ejemplo, el arseniuro de galio.

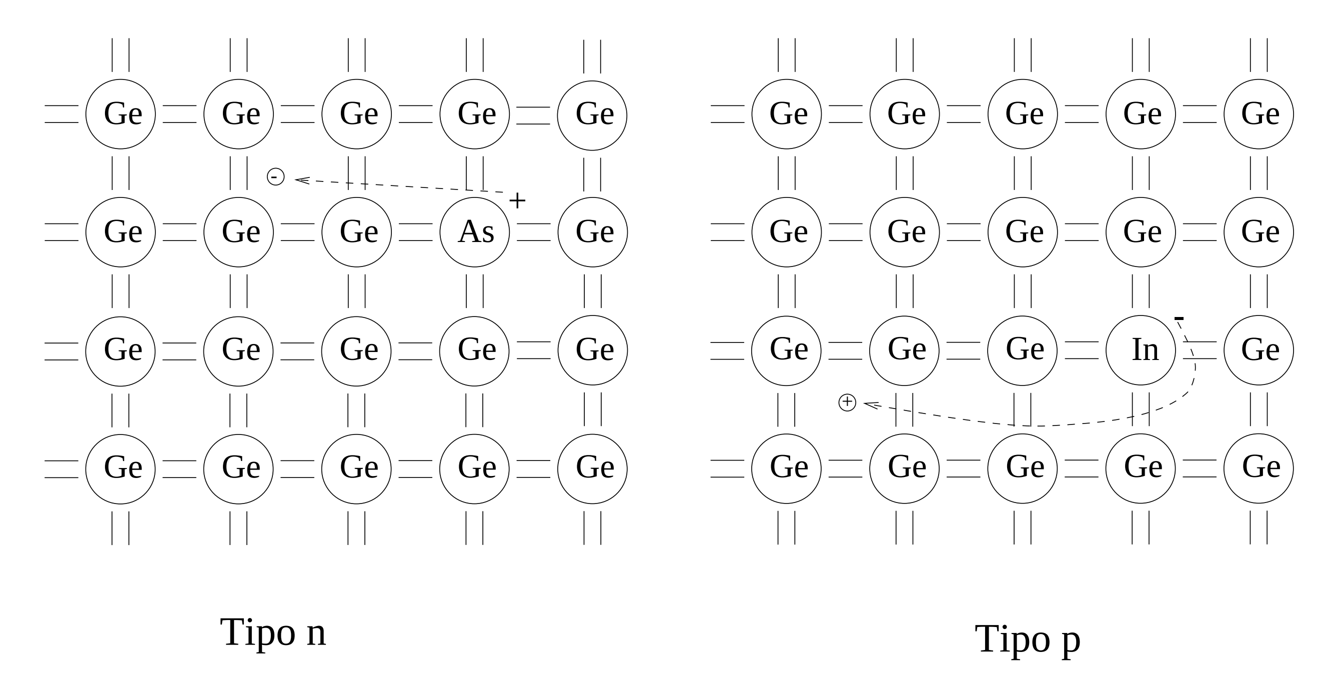


Figura 5.7: Electrones y huecos libres procedentes de átomos de impurezas (interpretación clásica).

Cuatro de ellos se usan para formar enlaces covalentes con átomos vecinos y el quinto se enlaza al átomo de impureza con fuerzas electrostáticas muy débiles. El enlace es tan débil que, incluso a temperaturas muy bajas (del orden de 20K), la energía de agitación térmica es suficientemente elevada como para ionizar al átomo de impureza [71], pasando el electrón a la banda de conducción sin producirse un hueco en la banda de valencia. Esto puede interpretarse como que existen estados energéticos accesibles para los electrones por debajo, pero muy próximos al mínimo de la banda de conducción, tal como se muestra en la fig. {5.6}. La proximidad es tan grande que el aporte de una cantidad mínima de energía (es decir, que la temperatura sea ligeramente superior a 0K) hace que estos electrones pasen a la banda de conducción. El átomo de impureza que ha perdido el electrón se convierte en un *ión positivo y fijo*. Esta situación se ilustra en la fig. {5.7}. Estos semiconductores se denominan de **tipo n**, designados así porque la mayoría de los portadores de carga son, a temperaturas normales de funcionamiento, electrones libres. La componente de conductividad eléctrica que se produce por los átomos de impureza se llama *conductividad por impureza*. Los átomos del grupo V de substitución se llaman con frecuencia *átomos donadores*, ya que cada uno de ellos dona un electrón libre al cristal.

Si en lugar de los átomos del grupo V se introducen en la red del semiconductor de tipo IV átomos del grupo III (Al,Ga,In,etc.), se observará un fenómeno muy distinto. Estos átomos tienen sólo tres electrones de valencia que se usan para formar enlaces covalentes con tres átomos vecinos, pero el cuarto posible enlace carece de electrón. Esto puede interpretarse como que existen estados energéticos accesibles para los electrones por encima, pero muy próximos al máximo de la banda de valencia, como se muestra en la fig. {5.6}. La cercanía es tan grande que el aporte de una cantidad mínima de energía (es decir, que la temperatura sea superior a unas cuantas decenas de grados Kelvin) hace que los electrones de la banda de valencia pasen a ocupar estos estados energéticos, produciéndose un hueco en la banda de valencia [71]. El átomo de impureza se convierte en un *ión negativo y fijo*. Esta situación se ilustra en la fig. {5.7}. Estos semiconductores se denominan de **tipo p**, designados así porque

la mayoría de los portadores de carga son, a temperaturas normales de funcionamiento, huecos libres. La componente de conductividad eléctrica que se produce por los átomos de impureza se llama *conductividad de impureza.* Los átomos del grupo III de substitución se llaman con frecuencia *átomos aceptores*, ya que cada uno de ellos acepta un electrón de la banda de valencia produciéndose un hueco adicional en la banda de valencia.

La estadística de ocupación de estados en los semiconductores tipo n y p se caracteriza por la presencia del nivel de Fermi por encima (para el tipo n) o por debajo (para el tipo p) del nivel de Fermi de un semiconductor puro (intrínseco). Según lo que vimos en el apartado anterior, en un cristal tipo n no puede haber más electrones que huecos a menos que el nivel de Fermi esté situado por encima del nivel de Fermi del semiconductor intrínseco. La situación contraria se da para un cristal tipo p.

En un semiconductor en equilibrio térmico debe haber un hueco o un ión positivo por cada electrón libre, y un electrón libre o un ión negativo por cada hueco libre (consideramos el caso genérico en el que haya simultáneamente impurezas donadoras y aceptoras). Por tanto, dado que el cristal debe ser eléctricamente neutro, esta condición de neutralidad de cargas se expresa igualando la suma algebraica de cargas positivas (fijas y móviles) y negativas (fijas y móviles):

$$p_\circ + N_d = n_\circ + N_a \tag{5.9}$$

donde $p_\circ$, $n_\circ$ son las concentraciones de huecos y electrones generados térmicamente y N_d, N_a las concentraciones de cargas fijas producidas por la ionización de las impurezas donadoras y aceptoras. Puede probarse que sustituyendo en esta expresión (5.5) y (5.6) se llega a:

$$E_f = \frac{1}{2}(E_v + E_c) + \frac{3}{4}kT\ln\frac{m_h}{m_e} + kT\sinh^{-1}\left(\frac{N_d - N_a}{2\sqrt{U_cU_v}e^{-\frac{\Delta E}{2kT}}}\right) \tag{5.10}$$

Los dos primeros sumandos representan el nivel de Fermi del semiconductor si éste fuese intrínseco, E_{fi}. Teniendo en cuenta (5.8) la fórmula anterior puede expresarse de la siguiente forma:

$$E_f = E_{fi} + kT\sinh^{-1}\left(\frac{N_d - N_a}{2n_i}\right) \tag{5.11}$$

Esta es la expresión del nivel de Fermi para un semiconductor con impurezas. En la fig. {5.8} se observa la posición relativa que ocupan los niveles de Fermi en un semiconductor de tipo n y en uno de tipo p frente al nivel intrínseco. Cuanto mayor sea el dopado neto del semiconductor, mayor será el desplazamiento. Es importante destacar que para deducir (5.11) se ha empleado la aproximación de Boltzmann. Si el dopado neto es muy elevado, el nivel de Fermi se aproxima mucho (puede, incluso, superarlo), en el primer caso, al mínimo de la banda de conducción y en el otro, al máximo de la banda de valencia. Sugerimos al lector que estime para qué nivel de dopado (5.11) deja de ser válida. Para estimarlo se puede tomar como criterio admitir que la aproximación de Boltzmann deja de ser cierta cuando el nivel de Fermi se aproxima a una de las bandas a una distancia inferior o igual a $3kT$.

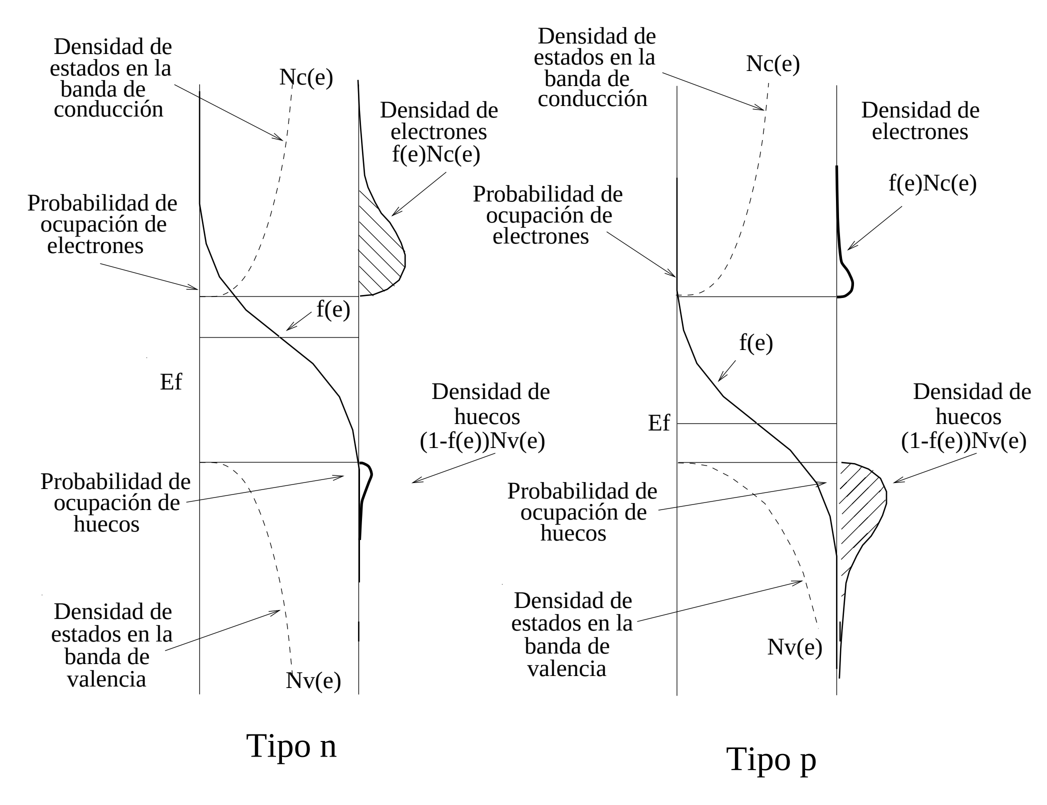


Figura 5.8: Electrones y huecos libres en semicondutores dopados.

La expresión (5.11) puede, en la mayoría de los casos, simplificarse. Normalmente el dopado en semiconductores es lo suficientemente elevado como para que la concentración de portadores mayoritarios procedente de la ionización de impurezas sea muy grande en comparación con la concentración de portadores mayoritarios de origen térmico. Esto quiere decir que el argumento de la función $\sinh^{-1}$ es muy grande. En estas circunstancias puede hacerse la aproximación: $\sinh^{-1} x \simeq \ln|2x|$, con lo que resulta:

$$E_f = E_{fi} \pm kT \ln\left(\frac{|N_d - N_a|}{n_i}\right) \tag{5.12}$$

en donde el signo más se utiliza para materiales de tipo n ($N_d > N_a$), y el signo menos para materiales de tipo p ($N_d < N_a$). El error cometido en la aproximación es del veinte por ciento cuando $n_i = N_d - N_a$ e inferior al cinco por ciento (decreciendo además rápidamente) para $n_i = 0{,}5(N_d - N_a)$. Aconsejamos al lector que utilice esta expresión para determinar cómo varía el nivel de Fermi de un semiconductor extrínseco con respecto a la temperatura.

La concentración de portadores en las bandas de conducción y valencia se puede encontrar acudiendo a la ecuación de neutralidad de cargas. Por ejemplo, cuando en (5.9) se sustituye $p_\circ = \frac{n_i^2}{n_\circ}$ se obtiene, resolviendo la ecuación de segundo grado resultante [2]:

$$n_\circ = \frac{1}{2}(N_d - N_a) + \sqrt{\frac{1}{4}(N_d - N_a)^2 + n_i^2} \tag{5.13}$$

Del mismo modo, cuando se sustituye $n_o = \frac{n_i^2}{p_\circ}$, se obtiene:

$$p_\circ = -\frac{1}{2}(N_d - N_a) - \sqrt{\frac{1}{4}(N_d - N_a)^2 + n_i^2} \tag{5.14}$$

De acuerdo con estas fórmulas, es evidente que cuando $N_d - N_a = 0$, $n_\circ = p_\circ = n_i$. Este es el caso de compensación total de dopados y el semiconductor se comporta como si fuera intrínseco. Si $N_d - N_a > 0$, $n_\circ > p_\circ$, como es de esperarse en un semiconductor de dopado tipo n; si $N_d - N_a < 0$, $n_\circ < p_\circ$, como se requiere en un semiconductor de tipo p.

Para semiconductores fuertemente extrínsecos (es el caso habitual), estas fórmulas se simplifican mucho. Por ejemplo, si $N_d - N_a \gg n_i$, como sucede para un semiconductor de fuerte dopado tipo n, resulta:

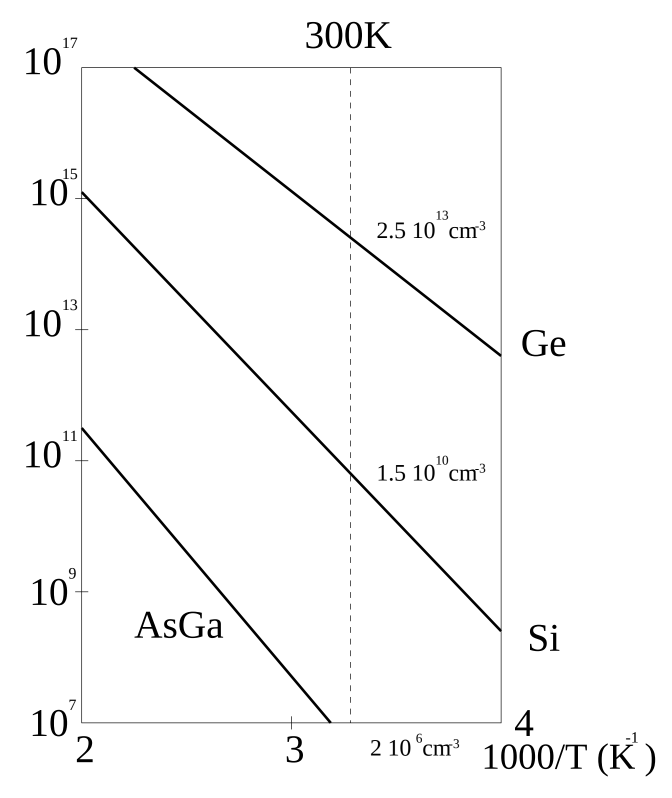


Figura 5.9: Concentración de portadores intrínsecos en función de la temperatura.

$$\begin{aligned} n_\circ &\simeq N_d - N_a \\ p_\circ &\simeq \frac{n_i^2}{N_d - N_a} \end{aligned} \tag{5.15}$$

en tanto que para un semiconductor tipo p muy dopado, en donde $N_d - N_a$ es grande y negativo,

$$\begin{aligned} p_\circ &\simeq N_a - N_d \\ n_\circ &\simeq \frac{n_i^2}{N_a - N_d} \end{aligned} \tag{5.16}$$

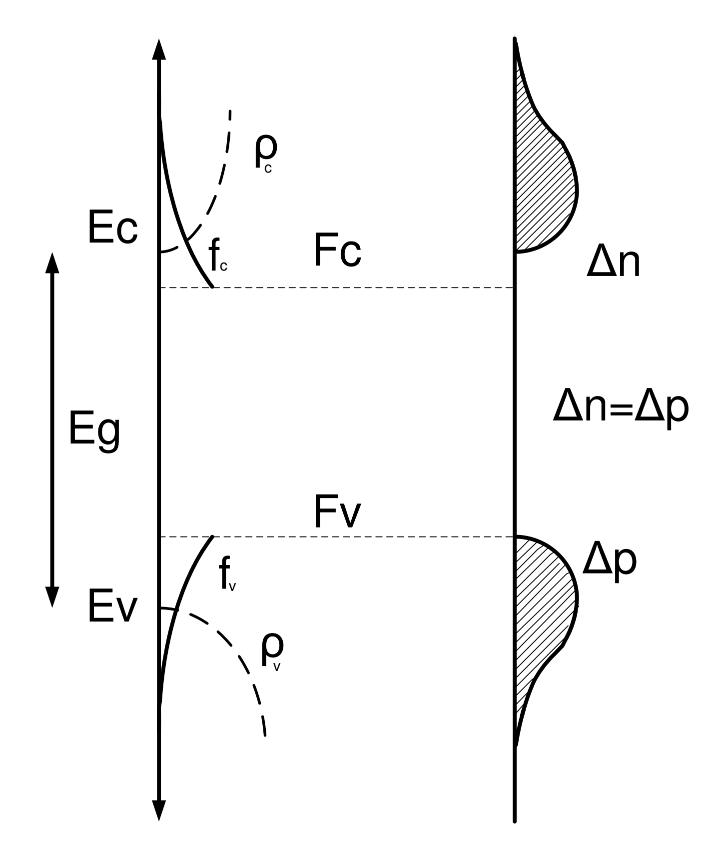


Figura 5.10: Concentración de portadores en un semiconductor fuera del equilibrio térmico.

Las expresiones anteriores parecen indicar que, mientras la concentración de portadores minoritarios depende fuertemente de la temperatura (n_i depende de la temperatura), la de los portadores mayoritarios no depende de ésta. Esta afirmación es cierta en un rango de temperaturas limitado. En (5.13) y (5.14) se ha despreciado la contribución de los portadores intrínsecos frente a los extrínsecos a la hora de determinar la concentración de portadores mayoritarios. Sin embargo, tal como se muestra en la fig. {5.9}, esta contribución crece rápidamente con la temperatura, por lo que cuando ésta es suficientemente elevada, la contribución de portadores intrínsecos se hace comparable a la de los extrínsecos [77]. Sugerimos al lector que estime a qué temperatura un semiconductor, que en principio era extrínseco, pasa a comportarse como uno intrínseco.

5.3. Semiconductores fuera del equilibrio térmico.

En un semiconductor en equilibrio térmico, los electrones de la banda de conducción están en equilibrio con los huecos de la banda de valencia debido a que ambas bandas intercambian energía gracias a los procesos de generación y recombinación de portadores [71]. Por ello existe una única función de distribuición que describe la probabilidad de ocupación de los estados energéticos accesibles en ambas bandas. Si ejercemos alguna acción externa sobre

el material (inyectamos una corriente, por ejemplo) esto deja de ser cierto. Los electrones de la banda de conducción están en una situación de cuasiequilibrio entre ellos (estos electrones son capaces de intercambiar energía rápidamente mediante colisiones con la red cristalina si su concentración es baja y por colisiones directas entre los propios portadores si su concentración es alta [20]). Lo mismo ocurre con los huecos de la banda de valencia, pero ambas bandas no están en equilibrio mutuo. Esto se refleja en la aparición de dos funciones de distribución distintas, una para los electrones en la banda de conducción y otra para los huecos en la banda de valencia. Estas dos funciones están perfectamente definidas si conocemos, respectivamente, la posición de los *cuasi niveles de Fermi* en la banda de conducción y en la banda de valencia. Lógicamente, la posición de los dos cuasiniveles de Fermi no es independiente la una de la otra. Sus posiciones determinan las concentraciones totales de electrones y huecos y, dado que el semiconductor tiene que ser eléctricamente neutro [6], se tiene que cumplir que $\Delta n = \Delta p$.

La manera de calcular la concentración de portadores en cada banda sigue siendo la misma: tenemos que multiplicar la densidad de estados energéticos accesibles por la probabilidad de que estén ocupados (si lo que estamos calculando es la concentración de electrones en la banda de conducción) y por la probabilidad de que estén vacíos (si lo que estamos calculando es la concentración de huecos en la banda de conducción). Si integramos para todos los valores energéticos de la banda de conducción tendremos la concentración total de electrones y si procedemos de igual forma en la banda de valencia obtendremos la concentración total de huecos:

$$\begin{aligned} dn(E) &= N_e(E) f(E, F_c, T) dE = \frac{8\sqrt{2}\pi}{h^3} m_e^{\frac{3}{2}} \sqrt{E - E_c} \frac{1}{1 + e^{\frac{(E-F_c)}{kT}}} dE \\ dp(E) &= N_h(E)[1 - f(E, F_v, T)] dE = \frac{8\sqrt{2}\pi}{h^3} m_h^{\frac{3}{2}} \sqrt{E_V - E} \frac{1}{1 + e^{\frac{(E-F_v)}{kT}}} e^{\frac{(E-F_v)}{kT}} dE \end{aligned} \quad (5.17)$$

Podemos expresar matemáticamente la relación que existe entre el cuasinivel de Fermi en la banda de conducción y el de la banda de valencia, mediante las siguientes relaciones integrales:

$$\begin{aligned} \int_{E_c}^{\infty} \frac{8\sqrt{2}\pi}{h^3} m_e^{\frac{3}{2}} \sqrt{E_2 - E_c} \frac{1}{1 + e^{\frac{(E_2-F_c)}{kT}}} dE &= n_\circ + \Delta n \\ \int_{-\infty}^{E_v} \frac{8\sqrt{2}\pi}{h^3} m_h^{\frac{3}{2}} \sqrt{E_V - E_1} \frac{1}{1 + e^{\frac{(E_1-F_v)}{kT}}} e^{\frac{(E_1-F_v)}{kT}} dE &= p_\circ + \Delta n \end{aligned} \quad (5.18)$$

Donde $n_\circ$ y $p_\circ$ son las concentraciones de los portadores en la situación de equilibrio térmico. Vemos, pues, que no se pueden fijar independientemente las posiciones de los dos cuasiniveles de Fermi. Dada la posición de uno, podemos calcular la del otro, por el sencillo motivo de que el semiconductor es eléctricamente neutro. Normalmente en el medio activo de un láser se conoce la concentración total de portadores en desequilibrio $\Delta n = \Delta p = N$, de manera que las relaciones (5.18) sirven para determinar, aplicando métodos numéricos, la posición de los cuasiniveles de Fermi F_c y F_v.

[6] En los "puntos cuánticos" que, entre otras muchas aplicaciones, se emplean para fabricar láseres semiconductores, la neutralidad de carga no se verifica necesariamente.

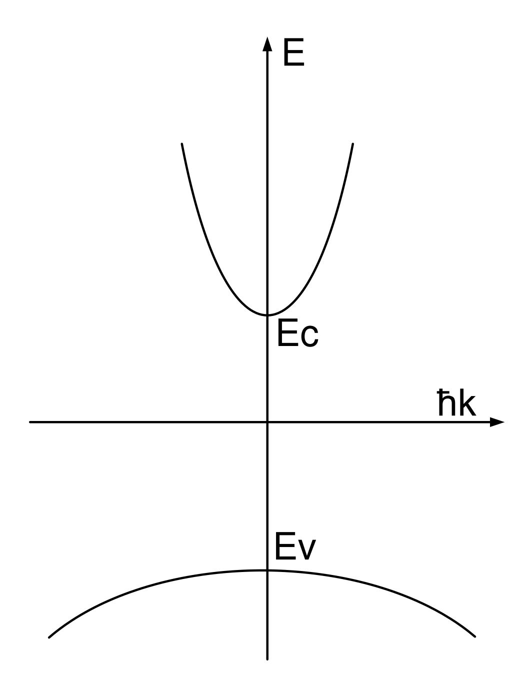


Figura 5.11: Diagrama de bandas de un semiconductor de transición directa.

5.4. Semiconductores de transición directa e indirecta.

Las funciones de onda asociadas a los electrones que se encuentran en estados estacionarios en un sólido cristalino, es decir, estados que no cambian con el tiempo, pueden determinarse de forma aproximada, ya sea aplicando teorías empíricas de pseudopotenciales o, si sólo interesan los estados próximos a los máximos y mínimos de las bandas de valencia y conducción, empleando el método $k.p$ de Kane [44]–[74]. En cualquier caso, estas funciones de onda representan estados electrónicos espacialmente no localizados dentro del cristal. Si un electrón se encuentra en un estado definido por una de estas funciones de onda, no se encuentra en una posición dada del cristal pero sí se conoce de una manera exacta la energía que tiene y el valor de una magnitud vectorial denominada *cuasiimpulso* o *impulso del cristal*, que está estrechamente relacionada con la cantidad de movimiento. Esto no debe extrañar. Recuerdemos que la primera relación de incertidumbre de Heisenberg afirma que $\Delta p \Delta r > \frac{\hbar}{2}$ por lo que si una partícula posee una cantidad de movimiento perfectamente determinada, su posición en el espacio es totalmente incierta. Esto es lo que ocurre en el cristal, un electrón descrito por una función de onda estacionaria posee un valor conocido del cuasiimpulso (magnitud relacionada con la cantidad de movimiento) pero su posición en el cristal está casi totalmente indeterminada. *El diagrama de bandas de un sólido relaciona las energías accesibles para los electrones en sus estados estacionarios dentro de una banda con los valores de cuasiimpulso de dichos electrones poseen.*

Un sólido conduce la electricidad porque los electrones se desplazan en su seno siguiendo una trayectoria determinada, por lo que ocupan una posición en el cristal y se desplazan con una velocidad determinada. Por consiguiente, las funciones de onda de los estados es-

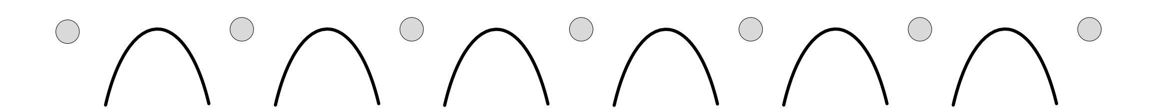


Figura 5.12: Potencial periódico en el interior de un cristal (representación unidimensional).

tacionarios del electrón en el sólido no describen el comportamiento de estos electrones. Sin embargo hallando combinaciones lineales (estas sumas decriben paquetes de ondas) de estas funciones puede describirse un electrón de este tipo.

El cuasiimpulso, al igual que la cantidad de movimiento, es una magnitud vectorial, por lo que, en principio, el diagrama de bandas exige representar una variable dependiente, la energía, en función de tres variables independientes, las componentes del cuasiimpulso.

En la fig. {5.11} se ha representado el diagrama de bandas simplificado del Arseniuro de Galio. El Arseniuro de Galio es un semiconductor de *transición directa*, es decir, el mínimo de energía de la banda de conducción y el máximo de la banda de valencia ocurren para un mismo valor del cuasiimpulso; además, para incrementos pequeños del cuasiimpulso, la energía no depende más que del módulo de éste [7], de ahí que sea posible hacer la representación bidimensional mostrada en la figura. Las curvas representadas son *parábolas*, al menos para valores pequeños de la variable independiente. Esto nos permite escribir la siguiente expresión para la banda de conducción (que refleja la relación cuadrática que existe entre la energía y el cuasiimpulso):

$$E_2 - E_c = \frac{\hbar^2 k^2}{2m_e} \tag{5.19}$$

donde E_c es la energía del mínimo de la banda de conducción, m_e es un término cuya inversa define la curvatura de la banda y que recibe el nombre de masa efectiva de los electrones en la banda de conducción y no coincide, en general, con la masa en reposo: vemos, por tanto, que el electrón se comporta en la banda de conducción como una partícula libre de masa m_e. Es posible dar una explicación intuitiva a esto. La relación que existe en Mecánica Clásica entre la energía y la cantidad de movimiento de una partícula libre es $E = \frac{p^2}{2m}$, donde m es la masa de la partícula. El equivalente clásico del cuasiimpulso es la cantidad de movimiento, de forma que, estableciendo una analogía entre esta expresión y (5.19), se concluye que m_e es una "masa efectiva". El motivo por el que m_e no coincide con la masa del electrón es que éste no es en absoluto una particula libre, sino que se encuentra sometido a un potencial periódico producido por los núcleos de la red cristalina [20]–[93], tal como se observa en la fig. {5.12}. Los portadores libres de carga en la banda de valencia (huecos) también podemos considerarlos como partículas libres, siempre que sustituyamos

[7] Esto no es rigurosamente cierto. La banda de valencia tiene tres subbandas distintas [20]–[93]: la de huecos pesados, la de huecos ligeros y otra subbanda adicional, de mayor energía, que no juega un papel importante. Las superficies de Fermi [12] de los denominados "huecos pesados" demuestran con claridad que su masa efectiva no es isotrópica.

su masa real, por una masa efectiva que denotaremos m_h. La relación que existe entre la energía de los huecos y el cuasiimpulso que poseen viene dada por la relación:

$$E_c - E_1 = \frac{\hbar^2 k^2}{2m_h} \tag{5.20}$$

Los semiconductores de transición directa reciben también el nombre de semiconductores ópticamente activos, porque un porcentaje importante de las recombinaciones son recombinaciones radiativas, es decir, la energía perdida por el electrón al pasar del mínimo de la banda de conducción al máximo de la banda de valencia se transforma en un fotón. Esto es así, porque las recombinaciones directas entre bandas son bastante probables, ya que no implican un cambio del cuasiimpulso.

En otros semiconductores, es el caso del Silicio y el Germanio, el mínimo de la banda de conducción y el máximo de la banda de valencia no coinciden, por lo que reciben el nombre de semiconductores de *transición indirecta*. El máximo de la banda de valencia ocurre en ambos para el valor $\hbar\vec{k}$=0, pero los mínimos absolutos (son varios) de la banda de conducción se dan para valores de $\hbar\vec{k}$ distintos de cero. Además, aunque el máximo de la banda de valencia de estos cristales tiene una curvatura (masa efectiva) que no depende de la dirección de $\vec{k}$, no ocurre lo mismo con los mínimos de la banda de conducción, por lo que la masa efectiva de los electrones en la banda de conducción depende de la dirrección (se trata, en realidad, de una magnitud tensorial).

Por motivos que no van a explicarse [12]–[62], para los semiconductores de transición indirecta también es posible definir una masa efectiva única en la banda de conducción y tratar a los electrones como partículas libres con una cierta masa efectiva m_e. Este tipo de semiconductores se llaman también semiconductores ópticamente inactivos, porque la mayor parte de las recombinaciones son recombinaciones no radiativas, es decir, la energía perdida por el electrón al pasar del mínimo de la banda de conducción al máximo de la banda de valencia no se transforma en un fotón. Esto es así, porque las recombinaciones directas entre bandas son muy improbables, ya que implican un cambio del cuasiimpulso que tiene que ser cedido o absorbido por la red cristalina. En estos semiconductores la mayor parte de las recombinaciones tienen lugar en "trampas" [71], que son estados energéticos situados en la banda prohibida que aparecen como consecuencia de imperfecciones en la red cristalina.

La inmensa mayoría de los dispositivos de estado sólido emisores de luz (diodos LED y láser) se construyen, lógicamente, con semiconductores de transición directa y, en la segunda y tercera ventanas de transmisión, son aleaciones cuaternarias derivadas del AsGa y del PIn (fosfuro de indio) [102].

5.5. Densidad de estados compatibles.

En el apartado anterior hemos visto que para que la recombinación de un par electrón-hueco sea radiativa, es necesario que el cuasiimpulso se conserve, motivo por el que los semiconductores de transición directa son ópticamente activos y los de transición indirecta no. Pero las consecuencias de este razonamiento las podemos extender a un semiconductor

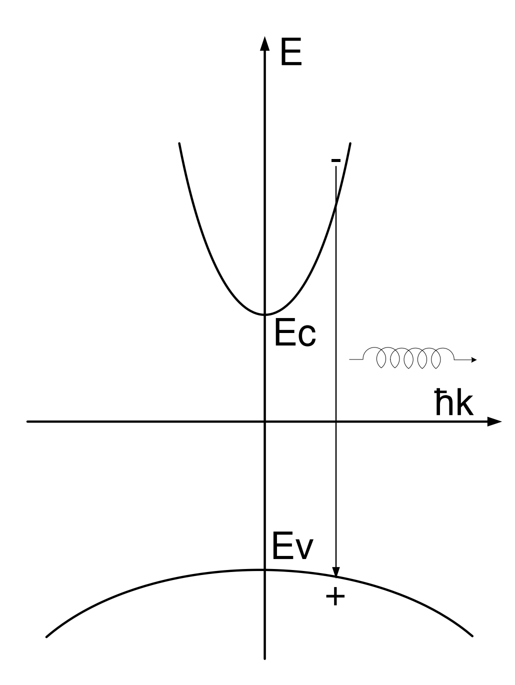


Figura 5.13: Recombinación radiativa de un par electrón-hueco en un semiconductor de transición vertical.

de transisión directa: un electrón de la banda de conducción sólo puede recombinarse con un hueco de la banda de valencia si ambos tienen el mismo valor de cuasiimpulso, es decir, si la transición es *vertical* en el diagrama de bandas [24], tal como se muestra en la fig. {5.13}. Imponiendo esta condición, podemos deducir la expresión de la densidad de estados energéticos compatibles. En este caso, "compatible" quiere decir que el paso del electrón de la banda de conducción a la banda de valencia, puede producir un fotón y viceversa, que un electrón de la banda de valencia pueda absorber un fotón y pasar a la banda de conducción a ocupar un estado energético libre. Para determinar esta densidad de estados compatibles partimos de la relación que hay entre la energía de un electrón en la banda de conducción y su cuasiimpulso y la relación análoga para un hueco en la banda de valencia:

$$\begin{aligned} E_2 &= E_c + \frac{\hbar^2 k^2}{2m_e} \\ E_v &= E_1 + \frac{\hbar^2 k^2}{2m_h} \end{aligned} \tag{5.21}$$

Como las transiciones tienen que ser verticales (el valor de $\hbar k$ en las igualdades anteriores es el mismo) y la energía de un fotón es $\hbar\omega = E_2 - E_1$, podemos sumar las dos ecuaciones anteriores y deducir:

$$\left.\begin{aligned} E_2 &= E_c + \frac{\hbar^2 k^2}{2m_e} \\ E_v &= E_1 + \frac{\hbar^2 k^2}{2m_h} \end{aligned}\right\} \Rightarrow \hbar\omega = E_2 - E_1 = E_g + \frac{\hbar^2 k^2}{2m_r} \Rightarrow \hbar^2 k^2 = 2m_r(\hbar\omega - E_g) \tag{5.22}$$

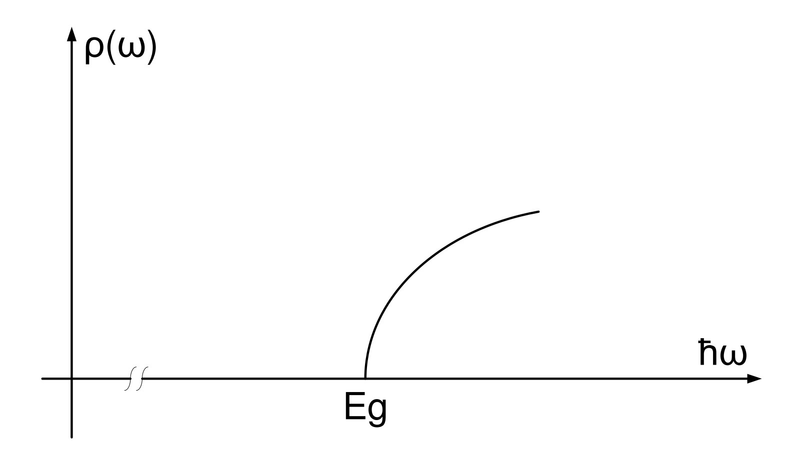


Figura 5.14: Densidad espectral de estados energéticos de un semiconductor que pueden interaccionar con la luz (por unidad de volumen).

donde: $\frac{1}{m_r} := \frac{1}{m_e} + \frac{1}{m_h}$ y $E_g := E_c - E_v$.

Pretendemos establecer la densidad de estados que pueden crear un fotón de energía $\hbar\omega$ o absorberlo (esta función va a intervenir en todos los procesos de interacción luz–semiconductor). Como existe una relación biunívoca entre los fotones y los estados energéticos de cualquiera de las dos bandas (por supuesto, imponiendo que las transiciones sean verticales), podemos escribir:

$$N_e(E_2)dE_2 = \rho(\omega)d\omega \tag{5.23}$$

y ahora las relaciones (5.21) y (5.22) nos permiten evaluar tanto $N_e(E_2)$ en función de ω como dE_2 en función de $d\omega$:

$$E_2 = E_c + \frac{m_r}{m_e}(\hbar\omega - E_g) \Rightarrow dE_2 = \frac{m_e}{m_c}\hbar d\omega \tag{5.24}$$

Lo único que resta hacer es evaluar la densidad de estados para los electrones de la banda de conducción (podríamos, igualmente, haber elegido los huecos en la banda de valencia y el resultado final sería el mismo) y tener en cuenta las relaciones que acabamos de deducir (que, en definitiva, resultan de imponer que la transición sea vertical en el diagrama de bandas):

$$N_e(E_2)dE_2 = \frac{8\sqrt{2}\pi}{h^3} m_e^{\frac{3}{2}} \sqrt{E_2 - E_c} dE_2 = \frac{(2m_e)^{3/2}}{2\pi^2\hbar^3}\left(\frac{m_r}{m_e}\right)^{1/2}\sqrt{\hbar\omega - E_g}\frac{m_r}{m_e}\hbar d\omega \Rightarrow$$

$$\rho(\omega) = \frac{(2m_r)^{3/2}}{2\pi^2\hbar^2}\sqrt{\hbar\omega - E_g} \tag{5.25}$$

5.6. Interacción luz-materia.

Pasamos ahora a estudiar los tres procesos de interacción luz-materia, con la particularidad de que el medio material es un sólido cristalino semiconductor.

5.6.1. Emisión espontánea.

Para que puedan producirse fotones por emisión espontánea de una determinada energía $\hbar\omega$ han de verificarse una serie de condiciones que pasamos a enumerar [87]:

1. Que existan estados energéticos en el semiconductor que puedan producir fotones de esa energía. Por tanto la tasa de emisión espontánea es proporcional a $\rho(\omega)$.
2. Que exista un electrón que se pueda recombinar. La tasa de emisión espontánea es, por tanto, proporcional a la probabilidad de que haya un electrón con la energía E_2, $f_c(E_2, F_c) = \frac{1}{1+e^{\frac{(E_2-F_c)}{kT}}}$.
3. Que exista un hueco que se pueda recombinar. La tasa de emisión espontánea es proporcional, pues, a la probabilidad de que haya un hueco con la energía E_1, $1-f_v(E_1, F_v) = \frac{e^{\frac{(E_1-F_v)}{kT}}}{1+e^{\frac{(E_1-F_v)}{kT}}}$.

Concluimos, entonces, que la tasa de emisión espontánea (número de fotones generados por unidad de volumen y unidad de tiempo en cada intervalo espectral) viene dada por la siguiente expresión:

$$r_{sp} = \gamma_{rad}\rho(\omega)f_c(E_2, F_c)[1 - f_v(E_1, F_v)] \tag{5.26}$$

En realidad las energías E_1 y E_2 están relacionadas con la energía del fotón ya que, tal como vimos en el apartado anterior, las transiciones entre bandas han de ser verticales, por lo que se verifica [25]:

$$\left.\begin{array}{l} E_2 = E_c + \dfrac{m_r}{m_e}(\hbar\omega - E_g) \\ \\ E_v = E_1 + \dfrac{mr}{m_h}(\hbar\omega - E_g) \end{array}\right\} \tag{5.27}$$

La ecuación (5.26) está sujeta a la restricción $\Delta n = \Delta p$ (que es una cantidad conocida si sabemos la corriente de polarización del láser), con lo que el único problema consiste en determinar la posición de los cuasiniveles de Fermi en función de la concentración de portadores. Para resolver este problema de forma exacta, habría que manipular numéricamente la función de Fermi–Dirac. Afortunadamente existen diversas relaciones aproximadas que

simplifican el problema y permiten calcular F_c y F_v en función de $\Delta n = \Delta p$. Damos, como ejemplo, las relaciones de Nilsson [25]:

$$\left.\begin{aligned} F_c &= \{\ln\delta + \delta[64 + 0{,}05524\delta(64 + \sqrt{\delta})]\} \\ F_v &= \{\ln\epsilon + \epsilon[64 + 0{,}05524\epsilon(64 + \sqrt{\epsilon})]\} \\ \delta &:= \frac{\Delta n}{2\left(\frac{m_e KT}{2\pi\hbar^2}\right)^{3/2}} \\ \epsilon &:= \frac{\Delta n}{2\left(\frac{m_h KT}{2\pi\hbar^2}\right)^{3/2}} \end{aligned}\right\} \tag{5.28}$$

Vamos a considerar ahora el caso particular de que el semiconductor no está excesivamente lejos del equilibrio térmico (es decir, la inyección de portadores es débil [8]), podemos hacer la aproximación de Boltzmann en las bandas de valencia y conducción, lo que nos lleva a:

$$\begin{aligned} f_c(E_2, F_c) &= \frac{1}{1+e^{\frac{(E_2-F_c)}{kT}}} \simeq e^{\frac{-(E_2-F_c)}{kT}} \\ 1 - f_v(E_1, F_v) &= \frac{e^{\frac{(E_1-F_v)}{kT}}}{1+e^{\frac{(E_1-F_v)}{kT}}} \simeq e^{\frac{(E_1-F_v)}{kT}} \end{aligned} \tag{5.29}$$

Si ahora sustituimos en la expresión de la tasa de emisión espontánea llegamos a:

$$r_{sp} = \frac{\gamma_{rad}(2m_r)^{3/2}}{\pi\hbar^2} e^{\frac{F_c-F_v-E_g}{kT}} \sqrt{\hbar\omega - E_g}\, e^{-\frac{\hbar\omega-E_g}{kT}} \tag{5.30}$$

Esta expresión predice la tasa de emisión espontánea de un semiconductor sólo en condiciones de inyección débil (es decir, incrementos en la concentración de portadores bajos en comparación con la concentración de portadores mayoritarios en condiciones de equilibrio térmico).

5.6.2. Emisión estimulada.

Para que puedan producirse fotones por emisión estimulada de una determinada energía $\hbar\omega$ se tienen que cumplir los siguientes requisitos [87]:

[8]En los láseres nunca se verifica esta hipótesis. La inyección de portadores siempre es fuerte ya que el cuasinivel de Fermi de los electrones siempre está situado dentro dela banda de conducción y el de los huecos cerca del máximo de la banda de valencia.

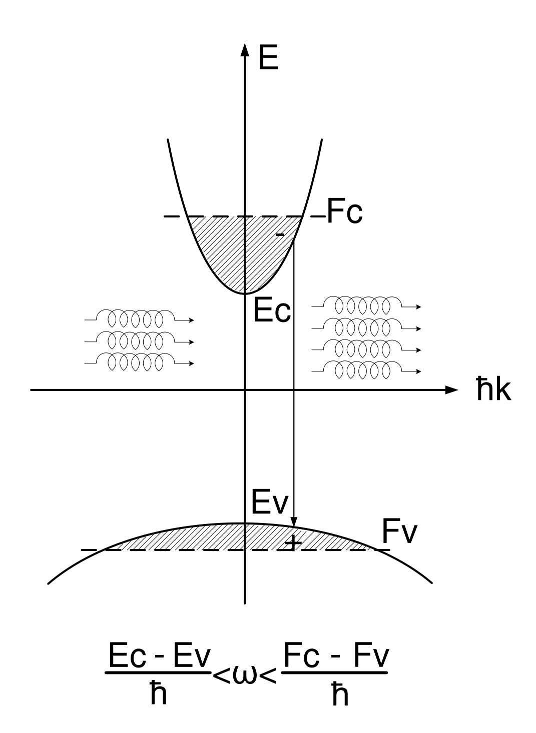


Figura 5.15: Intervalo de pulsaciones en las que el semiconductor puede amplificar.

1. Que existan estados energéticos en el semiconductor que puedan producir fotones de esa energía. Por tanto la tasa de emisión estimulada es proporcional a $\rho(\omega)$.

2. Que exista un electrón que se pueda recombinar de forma estimulada. La tasa de emisión estimulada es proporcional, por tanto, a la probabilidad de que haya un electrón con la energía E_2 en la banda de conducción, $f_c(E_2, F_c) = \frac{1}{1+e^{\frac{(E_2-F_c)}{kT}}}$.

3. Que exista un hueco con el que se pueda recombinar el electrón. La tasa de emisión estimulada es proporcional, por tanto, a la probabilidad de que haya un hueco con la energía E_1, $1 - f_v(E_1, F_v) = \frac{e^{\frac{(E_1-F_v)}{kT}}}{1+e^{\frac{(E_1-F_v)}{kT}}}$.

4. Que exista radiación (fotones) que estimulen a los electrones de la banda de conducción a recombinarse con huecos de la banda de valencia. Así, pues, la tasa de emisión de fotones estimulados es proporcional al flujo de fotones.

Por tanto la tasa de emisión estimulada (número de fotones generados por unidad de volumen y unidad de tiempo en cada intervalo espectral) viene dada por:

$$r_{est} = \Phi \frac{\lambda^2 \gamma_{rad}}{8\pi} \rho(\omega) f_c(E_2, F_c)[1 - f_v(E_1, F_v)] \qquad (5.31)$$

Repetimos, para reunir todas las expresiones necesarias para evaluar la tasa de emisión estimulada, las expresiones que relacionan E_1 y E_2 con $\hbar\omega$ y las relacions de Nilsson (dado

que suponemos que no vamos a tratar de invertir la función de Fermi–Dirac numéricamente) [25]:

$$\left.\begin{aligned} E_2 &= E_c + \frac{m_r}{m_e}(\hbar\omega - E_g) \\ E_v &= E_1 + \frac{mr}{m_h}(\hbar\omega - E_g) \\ F_c &= \{\ln\delta + \delta[64 + 0{,}05524\delta(64 + \sqrt{\delta})]\} \\ F_v &= \{\ln\epsilon + \epsilon[64 + 0{,}05524\epsilon(64 + \sqrt{\epsilon})]\} \\ \delta &:= \frac{\Delta n}{2\left(\frac{m_e KT}{2\pi\hbar^2}\right)^{3/2}} \\ \epsilon &:= \frac{\Delta n}{2\left(\frac{m_h KT}{2\pi\hbar^2}\right)^{3/2}} \end{aligned}\right\} \tag{5.32}$$

5.6.3. Absorción.

Para que puedan absorberse fotones de una determinada energía $\hbar\omega$ han de verificarse una serie de condiciones [87]:

1. Que existan estados energéticos en el semiconductor entre los que se pueda producir la absorción. Por tanto la tasa de absorción es proporcional a $\rho(\omega)$.

2. Que exista un electrón en la banda de valencia que pueda ser absorbido. La tasa de absorción es, pues, proporcional a la probabilidad de que haya un electrón con la energía E_1 en la banda de valencia, $f_v(E_1, F_v) = \frac{1}{1+e^{\frac{(E_1-F_v)}{kT}}}$.

3. Que exista un estado energético no ocupado en la banda de conducción. La tasa de absorción es proporcional, por tanto, a la probabilidad de que haya un estado libre en la banda de conducción de energía E_2, $1 - f_c(E_2, F_c) = \frac{e^{\frac{(E_2-F_c)}{kT}}}{1+e^{\frac{(E_2-F_c)}{kT}}}$.

4. Que exista radiación (fotones) que estimulen la absorción de electrones de la banda de valencia y los lleven a la banda de conducción. Por consiguiente, la tasa absorción de fotones es proporcional al flujo de éstos.

En definitiva, la tasa de absorción (número de fotones perdidos por unidad de volumen y unidad de tiempo en cada intervalo espectral) viene dada por la siguiente expresión:

$$r_{abs} = \Phi\frac{\lambda^2\gamma_{rad}}{8\pi}\rho(\omega)f_v(E_1, F_v)[1 - f_c(E_2, F_c)] \tag{5.33}$$

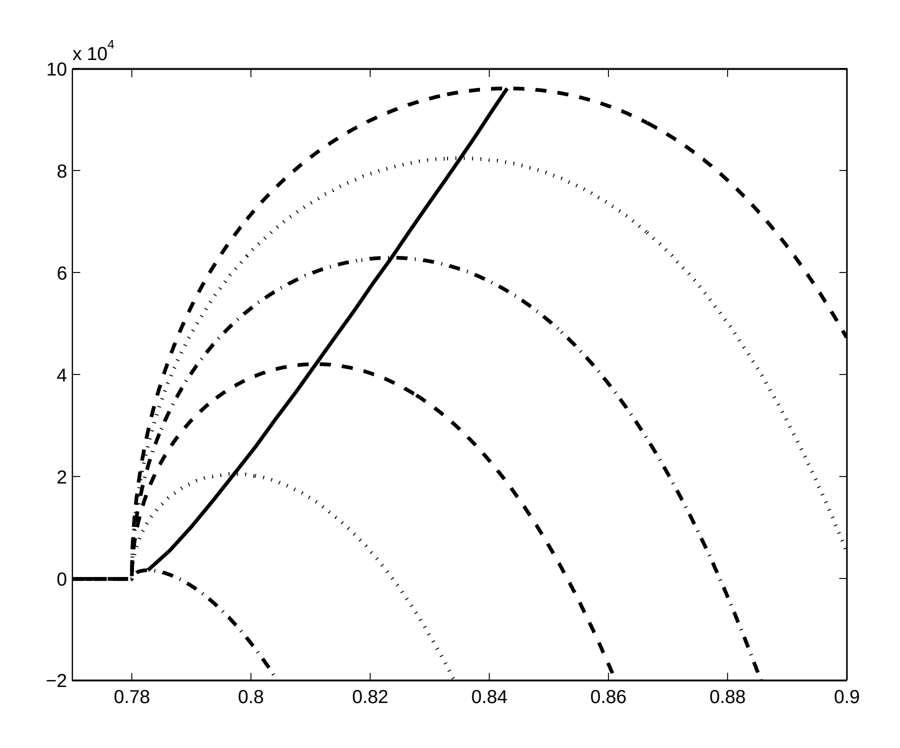


Figura 5.16: Evolución de la ganancia de un semiconductor (medida en m^{-1}) con respecto a la concentración de portadores y la energía del fotón (en eV). Las curvas representadas se corresponden con las siguientes concentraciones de portadores: $0,9\times10^{24}m^{-3}$, $1,34\times10^{24}m^{-3}$, $1,78\times10^{24}m^{-3}$, $2,22\times10^{24}m^{-3}$, $2,66\times10^{24}m^{-3}$ y $3\times10^{24}m^{-3}$. Simulación realizada por Alejandro Carballar Rincón.

que completamos con [25]:

$$\left.\begin{aligned}
E_2 &= E_c + \frac{m_r}{m_e}(\hbar\omega - E_g) \\
E_v &= E_1 + \frac{mr}{m_h}(\hbar\omega - E_g) \\
F_c &= \{\ln\delta + \delta[64 + 0{,}05524\delta(64+\sqrt{\delta})]\} \\
F_v &= \{\ln\epsilon + \epsilon[64 + 0{,}05524\epsilon(64+\sqrt{\epsilon})]\} \\
\delta &:= \frac{\Delta n}{2\left(\frac{m_e KT}{2\pi\hbar^2}\right)^{3/2}} \\
\epsilon &:= \frac{\Delta n}{2\left(\frac{m_h KT}{2\pi\hbar^2}\right)^{3/2}}
\end{aligned}\right\} \qquad (5.34)$$

5.7. Ganancia óptica del medio.

Para calcular el número neto de fotones cedidos por unidad de volumen y tiempo a la radiación que se propaga por el semiconductor (en cada intervalo espectral) tenemos que restarle a los que se ceden por emisión estimulada los que se pierden por absorción. Además sabemos que este incremento se traduce en una variación del flujo en la dirección de propagación [20]:

$$\begin{aligned} \frac{dN_f}{dt} &= \frac{\partial \Phi}{\partial z} = r_{est} - r_{abs} = \Phi \frac{\lambda^2 \gamma_{rad}}{8\pi} \rho(\omega)[f_c(E_2, F_c) - f_v(E_1, F_v)] \Rightarrow \\ \alpha(\omega) &:= \frac{1}{\Phi}\frac{\partial \Phi}{\partial z} = \frac{\lambda^2 \gamma_{rad}}{8\pi} \rho(\hbar\omega)[f_c(E_2, F_c) - f_v(E_1, F_v)] \end{aligned} \quad (5.35)$$

que, de nuevo, completamos con:

$$\left.\begin{aligned} E_2 &= E_c + \frac{m_r}{m_e}(\hbar\omega - E_g) \\ E_v &= E_1 + \frac{mr}{m_h}(\hbar\omega - E_g) \\ F_c &= \{\ln \delta + \delta[64 + 0{,}05524\delta(64 + \sqrt{\delta})]\} \\ F_v &= \{\ln \epsilon + \epsilon[64 + 0{,}05524\epsilon(64 + \sqrt{\epsilon})]\} \\ \delta &:= \frac{\Delta n}{2\left(\frac{m_e KT}{2\pi\hbar^2}\right)^{3/2}} \\ \epsilon &:= \frac{\Delta n}{2\left(\frac{m_h KT}{2\pi\hbar^2}\right)^{3/2}} \end{aligned}\right\} \quad (5.36)$$

Analizando la ecuación (5.35) podemos sacar algunas conclusiones cualitativas. En primer lugar, para que haya ganacia óptica, se tiene que cumplir que $f_c(E_2, F_c) - f_v(E_1, F_v) > 0$, lo que es equivalente a decir que la energía de los fotones debe ser inferior a la separación entre los cuasiniveles de Fermi y mayor que la energía de banda prohibida. El medio semiconductor es capaz de amplificar radiación electromagnética en el intervalo: $\frac{E_g}{\hbar} \leq \omega \leq \frac{F_c - F_v}{\hbar}$ [20], tal y como se muestra en la fig. {5.15}.

Ahora vamos a mostrar los resultados de simulación de la curva de ganancia óptica de un semiconductor que está a temperatura ambiente (T=300K). El semiconductor en cuestión

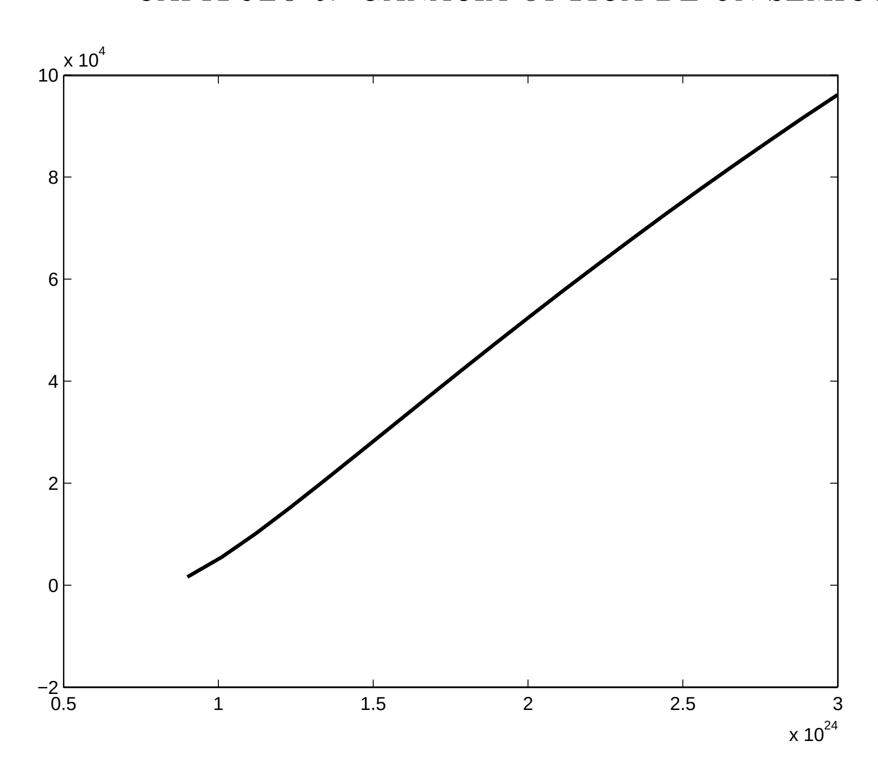


Figura 5.17: Ganancia máxima del semicondutor (medida en m^{-1}) frente a la concentración de portadores (medida en m^{-3}). Simulación realizada por Alejandro Carballar Rincón.

es una muestra de $InGaAsP$ con una anchura de banda prohibida $E_g = 0,78eV$, es decir, un semiconductor que emite luz a $\lambda_\circ = 1,55\mu m$. El tiempo medio de recombinaciones radiativas es $\tau_r = 0,9ns$ y las masas efectivas de electrones y huecos son $m_e = 0,046m$ y $m_h = 0,46m$ (donde m es la masa de un electrón libre). La curva está parametrizada para incrementos de la concentración de portadores cada vez mayores: $\Delta n = 0{,}9\times 10^{18} cm^{-3}$, $\Delta n = 1,34\times 10^{18} cm^{-3}$, $\Delta n = 1,78\times 10^{18} cm^{-3}$, $\Delta n = 2,22\times 10^{18} cm^{-3}$, $\Delta n = 2,66\times 10^{18} cm^{-3}$ y $\Delta n = 3\times 10^{18} cm^{-3}$. Esto se muestra en la fig. {5.16}.

Un láser emite luz en los modos longitudinales que están próximos a los máximos de las curvas que relacionan la ganacia frente a la energía de los fotones, puntos que se han marcados en la fig. {5.16}. Por tanto, lo que realmente nos interesa en este apartado, es relacionar los valores máximos de $\alpha(\omega)$ frente a la concentración de portadores $\Delta n = \Delta p$. La curva resultante se ajusta relativamente bien [20] por medio de una línea recta [9], tal como se muestra en la fig. {5.17}. Ésta es la relación que necesitamos para modelar el comportamiento de los láseres de materiales semiconductores. La expresión analítica de la ganacia es:

$$\alpha(\Delta n) = \alpha_{abs}(\frac{\Delta n}{\Delta n_T} - 1) = \frac{\alpha_{abs}}{\Delta n_T}(\Delta n - \Delta n_T) := g_N(N - N_\circ) \tag{5.37}$$

α_{abs} es el coeficiente de absorción del semiconductor, $\Delta n_T := N_\circ$ es la concentración de portadores de transparencia, es decir, la concentración de portadores en exceso (frente a la

[9] Esta relación es lineal sólo en los semiconductores volumétricos. En los pozos, líneas o puntos cuánticos esta dependencia no es lineal [20].

situación de equilibrio térmico) que debe haber para que el semiconductor sea transparente (es decir, que ni ampifique ni atenúe la luz que se propaga por el medio) y, por último g_n es la ganacia diferencial del medio.

5.8. Problemas

Problema 7: Se tiene un semiconductor intrínseco en el que se inyectan portadores de carga. La inyección es débil, lo que quiere decir que los cuasiniveles de Fermi no se alejan mucho de su posición de equilibrio térmico. Se trata de un semiconductor de transición directa por lo que va a emitir luz por emisión espontánea. Se desea conocer la pulsación a la que la densidad espectral de potencia óptica emitida es máxima.

Solución: La expresión de la tasa de emisión espontánea es:

$$r_{sp} = \frac{\gamma_{rad}(2m_r)^{3/2}}{\pi\hbar^2} e^{\frac{F_c - F_v - E_g}{kT}} \sqrt{\hbar\omega - E_g} e^{-\frac{\hbar\omega - E_g}{kT}}$$

La densidad espectral de potencia será máxima a aquella pulsación que verifique:

$$0 = \frac{\partial r_{sp}(\omega)}{\partial\omega} \simeq \frac{\gamma_{rad}(2m_r)^{3/2}}{\pi\hbar^2} e^{\frac{F_c - F_v - E_g}{kT}} \left[\frac{\hbar e^{-\frac{\hbar\omega - E_g}{kT}}}{2\sqrt{\hbar\omega - E_g}} - \frac{\hbar\sqrt{\hbar\omega - E_g}}{kT} e^{-\frac{\hbar\omega - E_g}{kT}} \right] \Rightarrow$$

$$0 = kT - 2(\hbar\omega_{max} - E_g) \Rightarrow \hbar\omega_{max} = E_g + \frac{kT}{2} \tag{5.38}$$

Por tanto la respuesta es:

$$\boxed{\hbar\omega_{max} = E_g + \frac{kT}{2}}$$

Capítulo 6

El láser de cavidad Fabry-Pérot

Resumen: En este capítulo vamos a estudiar el láser de cavidad Fabry-Pérot. En realidad vamos a establecer dos ecuaciones de balance, una para los portadores de carga dentro de la región activa y otra para los fotones dentro de la cavidad resonante. Estas dos ecuaciones, conocidas con el nombre de "rate–equations" en inglés, son válidas para cualquier tipo de láser, no sólo para los de cavidad FP, y permiten establecer la relación que hay entre la corriente inyectada y la evolución temporal de la potencia óptica total emitida. Estas mismas ecuaciones relacionan, también, la corriente inyectada con la variación en el tiempo de la concentración de portadores en la región activa. Además de plantear estas ecuaciones, explicaremos, cualitativamente, las características espectrales de estos láseres y describiremos su estructura física, es decir, veremos que la región activa de estos láseres es una doble heteroestructura en la que se consigue un confinamiento casi total de los portadores y un confinamiento parcial de la luz. Aunque la mayor parte de los láseres que se emplean hoy en día en comunicaciones ópticas son láseres de pozos cuánticos (y, quizás en el futuro lo sean de líneas o puntos quaánticos), estos láseres, cuyas regiones activas son semiconductores de bajas dimensiones (pozos cuánticos: dos, líneas cuánticas: una y puntos cuánticos: cero) no se abordan en este libro.

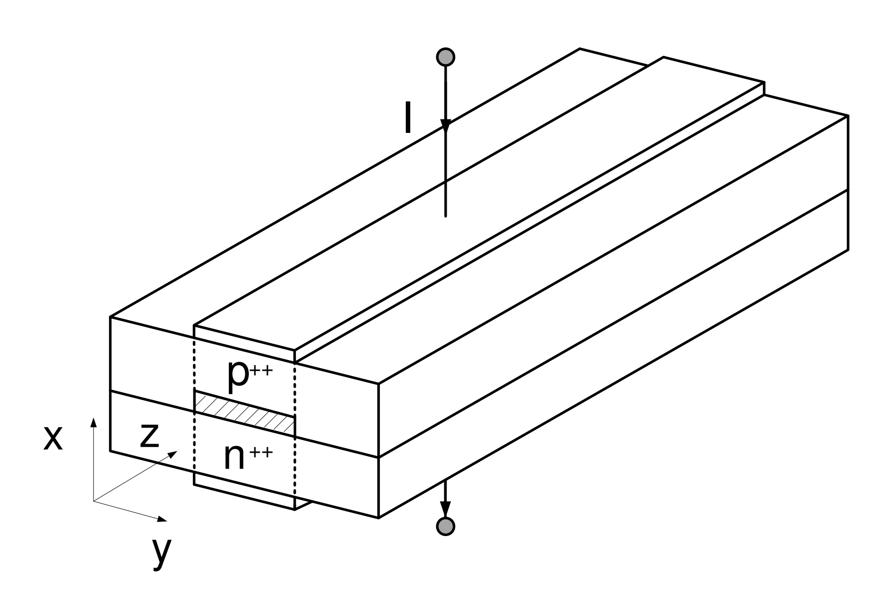


Figura 6.1: Estructura tridimendional de un diodo de cavidad Fabry-Pérot.

6.1. Introducción

En la fig. {6.1} se ha representado la estructura de un láser de cavidad Fabry-Pérot guiado por el índice. La capa de material semiconductor que se ha dibujado a rayas, recibe el nombre de *región activa* del láser y está constituida por un semiconductor ópticamente activo a la longitud de onda de trabajo (1300nm o 1550nm, normalmente) que está muy débimente dopado o que incluso puede ser intrínseco. El resto de semiconductores que rodean a la región activa, tanto en la dirección "x" como en la dirección "y" tienen unas anchuras de bandas prohibidas *mayores* que la de ésta. No se ha intentado representar los semiconductores que hay alrededor de la región activa, ya que son estructuras de geometrías muy complejas y que varían mucho de un láser a otro. En la bibliografía adjunta [4]–[102] se pueden encontrar varios ejemplos. El hecho de que la región activa tenga una anchura de banda prohibida menor que los semiconductores que hay alrededor tiene una importancia primordial para el correcto funcionamiento del dispositivo, ya que de esta manera se consiguen los siguientos objetivos [10]–[63]:

1. Un confinamiento *casi total* de los portadores en la región activa del láser.

2. Un confinamiento *parcial* de la luz en la región activa del láser.

Vamos a explicar detenidamente cada una de estas afirmaciones. Consideremos, primero, el diagrama de bandas de los semiconductores que aparecen en el plano $y = 0$, estando el diodo directamente polarizado, diagrama que se muestra en la fig. {6.2}. Supongamos que circula una corriente I por el dispositivo. La barrera de potencial que aparece en la sección "A" impide que los huecos se difundan por la zona n^{++} y quedan atrapados en la

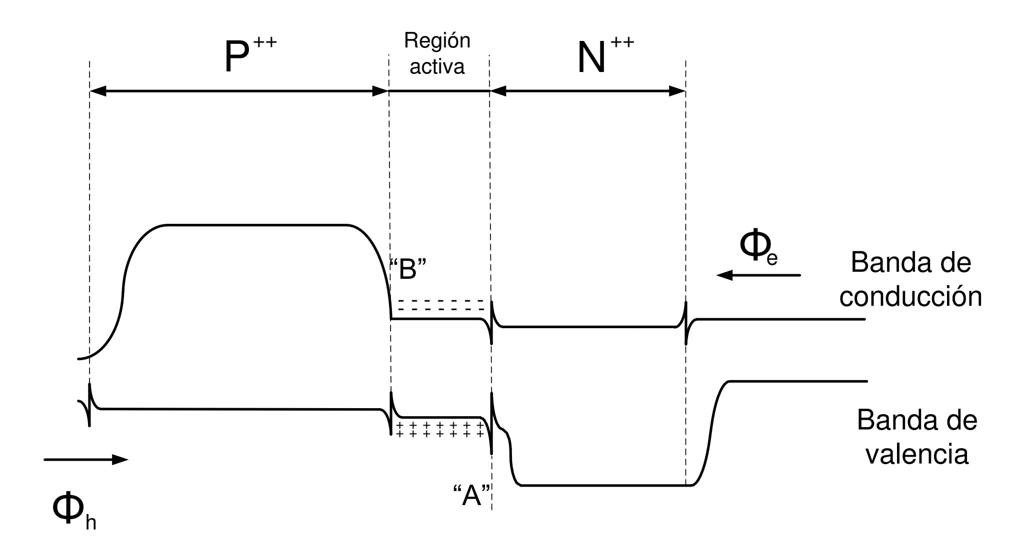


Figura 6.2: Diagrama de bandas de la doble heteroestructura de un diodo de cavidad Fabry-Pérot.

región activa. De igual forma, la barrera de potencial que aparece en la sección "B" en la banda de conducción, impide que los electrones se desplacen por difusión por la zona p^{++} y también quedan atrapados en la región activa. Por supuesto, en esta zona se produce una importante recombinación de pares electrón–hueco, que es la responsable de la corriente que está atravesando el diodo [1]. Vemos, pues, que la corriente $I = -qA_t\Phi_e = qA_t\Phi_h$ produce una acumulación de portadores en la región activa $\Delta n = \Delta p = N$ (ambas concentraciones son iguales, para que el semiconductor sea eléctricamente neutro [2]). Si esta concentración de portadores es suficientemente alta, la región activa tendrá un coeficiente de ganancia por unidad de longitud $\alpha > 0$, en vez de un coeficiente de atenuación, que es la situación que se da en equilibrio térmico y para concentraciones de portadores inferiores a un determinado valor, llamada concentración de transparencia que se denota $N_\circ$. En resumen, utilizando semiconductores con distintas anchuras de banda prohibida (las llamadas *dobles heteroestructuras*), conseguimos que el confinamiento de electrones y huecos se produzca en una determinada región del dispositivo llamada *región activa* [3].

[1]Estamos despreciando los portadores que saltan las barrera por efecto termoiónico y las corrientes difusivas que no llegan a la región activa. En un modelo más exacto del láser deberíamos cuantificar e incluir estos términos.

[2]Si no fuese así, se generarían unas fuerzas eléctricas muy intensas que consiguen, en un intervalo de tiempo muy corto, reestablecer el equilibrio de cargas. La hipótesis de neutralidad de cargas se aplica en el modelado de casi todos los dispositivos electrónicos y fotónicos.

[3]La idea de utilizar dobles heteroestructuras para conseguir un confinamiento simultáneo de portadores y fotones es, posiblemente, el hito más importante en la historia de los láseres a semiconductor, ya que hizo posible el funcionamiento de estos dispositivos a temperatura ambiente y en régimen continuo. La idea original la propusieron simultáneamente en 1963 H. Kroemer en EE.UU. y R.F. Kazarinov y Z.I. Alfero en la Unión Soviética .

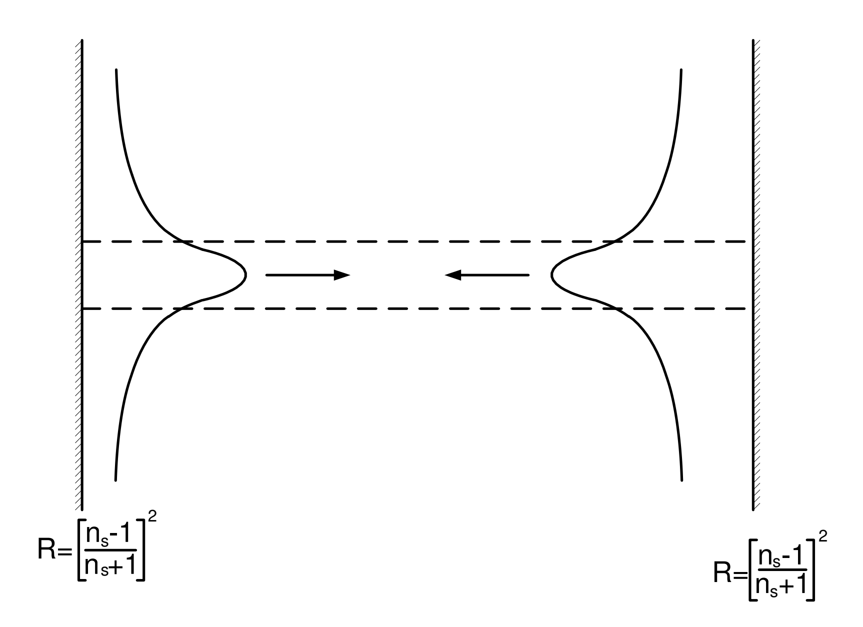


Figura 6.3: Distribución transversal de los modos guiados de una cavidad Fabry-Pérot guiada por el índice.

Los materiales semiconductores tienen una propiedad importante [1]: cuanto *menor* es su anchura de banda prohibida, mayor es su índice de refracción, por lo que la región activa tiene un índice de refracción mayor que el de todos los materiales que la rodean, haciéndose posible, por tanto, el guiado longitudinal de la luz. Recordemos que en las fibras ópticas la existencia de una región central (el núcleo) con un índice de refracción más alto que las regiones más alejadas del eje (la corteza) era lo que permitía que la luz se propagara longitudinalmente. El conjunto formado por la región activa y el resto de materiales semiconductores que la rodean constituyen, pues, un guiaonda que permite que la luz se propague longitudinalmente.

Tenemos, pues, una estructura de guiado longitudinal de la luz, en la que la región activa hace la función que en la fibra realiza el núcleo. Todos los materiales semiconductores que hay alrededor de ésta tienen anchuras de banda prohibida mayores, por lo que sus índices de refracciónson menores y cumplen una función análoga a la de la corteza en una fibra óptica. Las superficies frontal y posterior del diodo láser son cortes precisos a lo largo de planos cristalográficos del semiconductor, por lo que aparece una variación brusca del índice de refracción. En el Al_xAsGa_{1-x} el índice de refracción toma valores cercanos a 3,5, mientras que el del aire es prácticamente igual a la unidad, con lo que que estas superficies tienen un coeficiente de reflexión en potencia cercano la 31 % [4]. Así, pues, hay dos espejos planos paralelos entre sí situados en los extremos de la cavidad, es decir, tenemos una cavidad resonante.

El láser de cavidad Fabry Pérot se diseña de manera que a la longitud de onda de trabajo, exista un único modo transversal soportado por la estructura, tal como se muestra

[4]Los coeficientes de reflexión toman el valor $R = (\frac{n_s-1}{n_s+1})^2$.

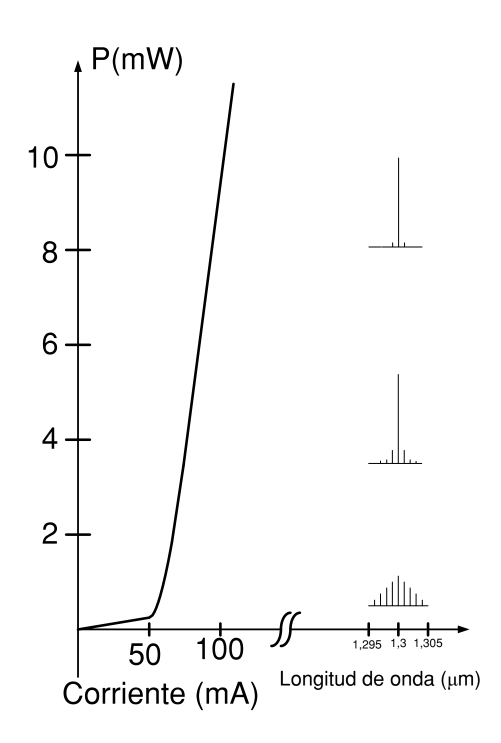


Figura 6.4: Espectro de un diodo de cavidad Fabry-Pérot guiado por el índice en función de la potencia emitida.

en la fig. {6.3}. Esto quiere decir que existe una única distribución transversal del campo electromagnético para la que es posible la propagación longitudinal de la luz láser. En esta figura su muesta también que sólo fracción de intensidad luminosa que se propaga por la cavidad solapa con la región activa y, por tanto, sólo esta fracción puede producir fenómenos estimulados. Esta fracción, expresada en tanto por uno, la denotamos por medio de Γ y recibe el nombre de factor de confinamiento.

El dispositivo emite a más de una longitud de onda, es decir, emite en varios modos longitudinales [4]. Los láseres guiados por el índice [4] (que son los que estamos considerando) tiene la propiedad de que cuanto mayor es la potencia de emisión y bajo condiciones de régimen permanente, son más monocromáticos, es decir, la distribución de potencia entre los distintos modos guiados se vuelve más asimétrica, de manera que a potencias altas, la mayor parte de ésta se emite a la longitud de onda correspondiente al modo longitudinal principal, tal y como se muestra en la fig. {6.4}.

En este apartado hemos identificado los elementos básicos del láser de cavidad Fabry-Pérot guiado por el índice (existen otros tipos de láseres de este tipo, los guiados por ganancia, que no vamos a considerar):

1. El medio amplificador, es decir, la región activa. Es un paralelepípedo compuesto de un material semiconductor de anchura de banda prohibida baja (en comparación con el resto de semiconductores que lo rodean) que es ópticamente activo a la longitud de onda de emisión del láser y que no está dopado.

2. El mecanismo de bombeo, que consiste en la inyección de una corriente, hace que la concentración de portadores en la región activa crezca y cuando ésta aumenta por

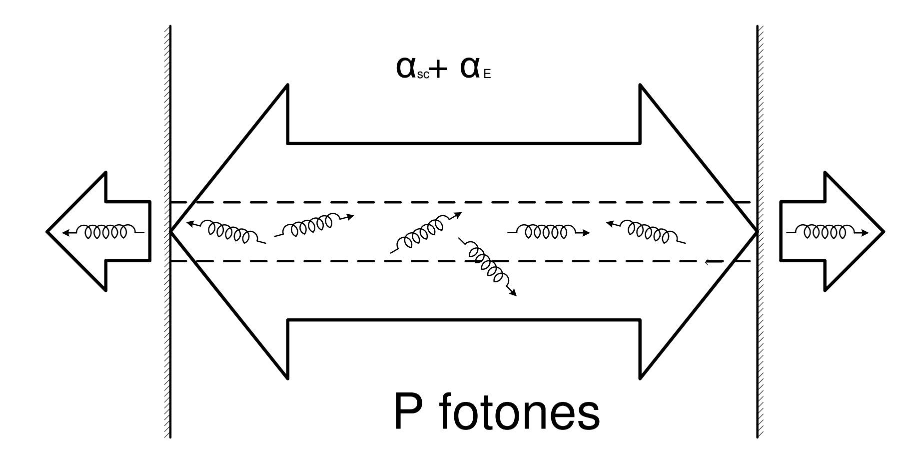


Figura 6.5: Cavidad resonante en la que hay radiación espontánea y estimulada.

encima de la concentración de transparencia, el medio activo adquiere una ganancia que es proporcional a la diferencia entre la concentración de portadores y la concentración de transparencia.

3. La cavidad resonante, que ocupa un volumen mayor que la región activa, ya que la distribución transversal del campo electromagnético se extiende mucho más allá de los límites de ésta. La cavidad resonante está limitada por dos planos cristalográficos paralelos entre sí, que tienen un coeficiente de reflexión en potencia cercano al 31 %.

En realidad, la mayoría de los láseres que se emplean hoy en día en comunicaciones ópticas son dispositivos de pozos cuánticos [109] y, quizás, en un futuro no muy lejano, sean láseres con medios activos constituidos por líneas y puntos cuánticos [44]–[74]–[97]. No vamos a abordar el estudio de estos emisores de luz porque no le presuponemos al lector los conocimientos de mecánica cuántica y física de semiconductores que son imprescindibles para tratar estas cuestiones.

6.2. Comportamiento dinámico de un láser

En este apartado vamos a deducir dos ecuaciones diferenciales no lineales acopladas entre sí que permiten determinar la evolución temporal de la potencia emitida por el láser en función de la corriente inyectada. Estas dos relaciones son ecuaciones de *continuidad*, que expresan la variación en el tiempo del número de portadores (o su concentración, si se prefiere) que hay en la *región activa* y la variación temporal del número de fotones (o su concentración) en la *cavidad resonante*. Es importante recalcar que estos dos volúmenes de control en los que vamos a deducir estas dos ecuaciones son distintos. La región activa es el parelelepípedo rayado representado en la fig. {6.1}, mientras que la cavidad resonante es el volumen completo del diodo semiconductor. Ambas son ecuaciones de balance, en las que

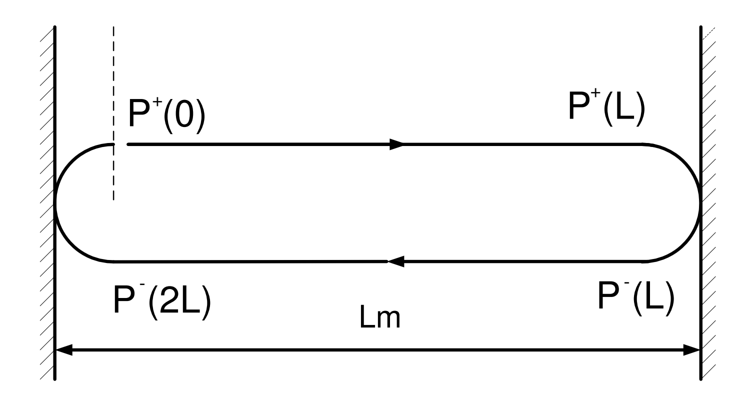


Figura 6.6: Cálculo de la constante de pérdidas en la cavidad resonante.

se evalúa la variación del número de fotones o portadores en su correspondiente volumen de control, como la suma de todos los términos que tienden a introducir fotones o portadores menos la suma de todos los que tienden a sacarlos fuera de allí.

Para poder plantear la ecuación de balance del número de fotones, tenemos que cuantificar los procesos que tienden a introducir fotones dentro de la cavidad y los que tienden a sacarlos. Consideremos primero el coeficiente de ganancia por unidad de longitud del medio activo. Evidentemente la ganancia del medio amplificador tiende a aumentar la concentración de fotones en su volumen de control. Deduzcamos la expresión exacta:

$$P = P_\circ e^{\Gamma \alpha_g x} \Rightarrow \left(\frac{dP}{dt}\right)_{est-abs} = \frac{dP}{dx}\frac{dx}{dt} = \Gamma \alpha_g P_\circ e^{\Gamma \alpha_g x}\frac{c_\circ}{n} := \frac{\Gamma P}{\tau_g} \tag{6.1}$$

En la expresión (6.1) hemos tenido en cuenta que sólo una fracción Γ (recordemos que Γ recibe el nombre de factor de confinamiento) de la intensidad luminosa total viaja por la región activa. Además hemos definido el tiempo medio que tarda en ganarse un fotón como resultado de los procesos estimulados (emisión estimulada menos absorción) como $\tau_g := \frac{n}{\alpha_g c_\circ}$. Si ahora tenemos en cuenta que $\alpha_g = g_n(N - N_\circ)$ resulta que $\tau_g = \frac{n}{g_n(N-N_\circ)c_\circ}$.

Ahora vamos a considerar los procesos físicos que sacan fotones de la cavidad resonante: el scattering Rayleigh y la reflexión parcial en los espejos. Para ello vamos a ver cómo disminuye el número de fotones como consecuencia de un viajes de ida y vuelta por la cavidad, tal como se muestra en la fig. {6.6}:

$$\begin{cases} P^+(0) = P_\circ \\ P^+(L) = P_\circ e^{-\alpha_{sc}L} \\ P^-(2) = RP_\circ e^{-\alpha_{sc}L} \\ P^-(2L) = RP_\circ e^{-\alpha_{sc}2L} \\ P^+(2L) = R^2 P_\circ e^{-\alpha_{sc}2L} \end{cases} \tag{6.2}$$

Vemos que el número de fotones va disminuyendo por culpa de las pérdidas por scattering Rayleigh y las reflexiones parciales en los espejos. Resulta conveniente repartir las pérdidas localizadas en los espejos mediante un coeficiente de pérdidas por unidad de longitud asociado a los espejos e introducirlo en el argumento de la función exponencial. Definiendo $\alpha_{esp} := \frac{1}{2L} Ln(\frac{1}{R^2})$, nos queda:

$$P^+(2L) = P_\circ e^{-(\alpha_{sc}+\alpha_{esp})2L} \Rightarrow P(x) = P_\circ e^{-(\alpha_{sc}+\alpha_{esp})x} \Rightarrow$$
$$\left(\frac{dP}{dt}\right)_{per} = \frac{dP}{dx}\frac{dx}{dt} = -2(\alpha_{sc} + \alpha_{esp})P\frac{c_\circ}{n} = \frac{P}{\tau_p} \tag{6.3}$$

donde la constante de tiempo de pérdidas de la cavidad fría (pérdidas que no están relacionadas con la atenuación del medio material) tiene la expresión: $\tau_p := \frac{n}{2c_\circ(\alpha_{sc}+\alpha_{esp})}$ y donde $\alpha_{esp} := \frac{1}{2L} Ln(\frac{1}{R^2})$. El significado físico de τ_p es evidente: es el tiempo que, en promedio, tarda en salir un fotón de la cavidad por alguno de los dos fenómenos antes mencionados, que son o bien una reflexión parcial en uno de los espejos, o bien un proceso de scattering Rayleigh.

Hay un último término que tiene tendencia a aumentar el número de fotones que hay en la cavidad: la pequeña fracción de fotones β del total generado por emisión espontánea que se propagan en la dirección perpendicular a los espejos y que cae dentro del ancho de banda de la luz generada por el láser. Si el número total de fotones generados por emisión espontánea por unidad de volumen y tiempo es $\frac{\Gamma N}{\tau_{esp}}$ los que se suman a la radiación generada por el láser es $\beta\frac{\Gamma N}{\tau_{esp}}$. Realmente no es necesario introducir este coeficiente empírico β [4]–[24], tal como veremos en el capítulo 10 donde se demostraremos que la emisión espntánea y la estimulada pueden relacionarse entre sí. La expresión que describe de forma aproximada la contribución de la emisión espontánea es:

$$\left(\frac{dP}{dt}\right)_{esp} = \beta\frac{\Gamma N}{\tau_{esp}} \tag{6.4}$$

En realidad el número de fotones generados por emisión espontánea por unidad de volumen y tiempo es, bajo condiciones de inyección fuerte (que son las habituales en la región activa de un láser), aproximadamente proporcional al cuadrado del número de portadores [5].

[5]En un semiconductor se producen tres tipos de recombinaciones [4]–[74]–[90]: recombinaciones en trampas, recombinaciones radiativas y recombinaciones Auger. Sólo las primeras permiten definir un tiempo de recombinación constante. Ni las recombinaciones en trampas ni las Auger contribuyen a la generación de fotones espontáneos.

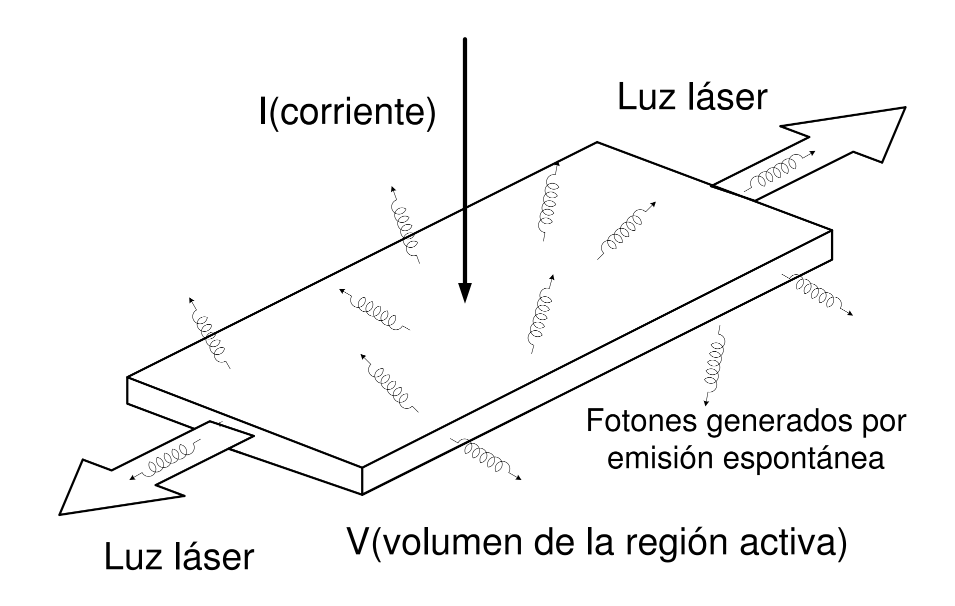


Figura 6.7: Volumen de control donde vamos a plantear la ecuación de continuidad de los portadores de carga.

La fórmula anterior, que es habitual en los modelos más sencillos del láser a semiconductor, la podemos aceptar como válida si suponemos que τ_{esp} depende de N, aunque, con objeto de simplificar el modelo, consideraremos que τ_{esp} es constante. Para modelos más realistas remitimos al lector a la bibliografía adjunta [20]–[24]–[25]–[74]. El coeficiente Γ aparece porque el volumen de la cavidad resonante es mayor que el de la región activa y el cociente de éste último dividido por el anterior es, precisamente, el factor de confinamiento. Ya hemos calculado todos los términos que tienden a cambiar la concentración de fotones dentro de la cavidad. Por tanto, teniendo en cuenta las expresiones (6.1), (6.3) y (6.4), podemos escribir la primera acuación de balance, la de los fotones dentro de la cavidad resonante:

$$\frac{dP}{dt} = \frac{\Gamma P}{\tau_g} - \frac{P}{\tau_p} + \beta\frac{\Gamma N}{\tau_{esp}}$$
$$\tau_g := \frac{n}{c_\circ g_m(N - N_\circ)}; \quad \tau_p = \frac{n}{2c_\circ(\alpha_{sc} + \alpha_{esp})} \tag{6.5}$$
$$\alpha_{esp} := \frac{1}{2L}\ln\left(\frac{1}{R^2}\right)$$

Vamos a plantear ahora la ecuación de continuidad del número de portadores en desequilibrio N (no especificamos si son electrones o huecos, ya que las concentraciones de desequilibrio tienen que ser iguales para que el semiconductor sea eléctricamente neutro) en el volumen de control definido por la región activa del diodo. Este volumen de control se muestra en la fig. {6.7}. El único término que contribuye a que la concentración de portadores en la región activa aumente es la inyección de corriente I y el aumento provocado es:

$$\left(\frac{dN}{dt}\right)_{corr} = \frac{I}{qV} \tag{6.6}$$

Para simplificar la ecuación de balance, suponemos que todos los portadores inyectados llegan a la región activa (lo cual no es cierto) y que los portadores no pueden saltar las barreras de potencial (una cierta proporción de ellos lo hace por efecto termoiónico). Los términos que tienden a hacer disminuir la concentración de portadores son el efecto neto emisión estimulada menos absorción y la tasa de emisión espontánea [6] (en este caso cuenta *todos* los fotones producidos por emisión espontánea, ya que todos aparecen como consecuencia de la recombinación de un pares electrón–hueco):

$$\begin{aligned} \left(\frac{dN}{dt}\right)_{est-abs} &= -\frac{P}{\tau_g} \\ \left(\frac{dN}{dt}\right)_{esp} &= -\frac{N}{\tau_{esp}} \end{aligned} \tag{6.7}$$

Ahora no aparece Γ en la primera expresión debido a que el volumen de la región activa es menor que el de la cavidad resonante. Por tanto, el par de ecuaciones de balance que describe, de forma simplificada, el comportamiento del láser es:

$$\left.\begin{aligned} \frac{dP}{dt} &= \frac{P\Gamma}{\tau_g} - \frac{P}{\tau_p} + \beta\frac{\Gamma N}{\tau_{esp}} \\ \frac{dN}{dt} &= \frac{I}{qV} - \frac{P}{\tau_g} - \frac{N}{\tau_{esp}} \\ \tau_g := \frac{n}{c_\circ g_m(N-N_\circ)};\quad \tau_p &= \frac{n}{2c_\circ(\alpha_{sc}+\alpha_{esp})} \\ \alpha_{esp} &:= \tfrac{1}{2L}Ln(\tfrac{1}{R^2}) \end{aligned}\right\} \tag{6.8}$$

El modelo dinámico del láser viene, pues, descrito por un par de ecuaciones diferenciales *no lineales* y acopladas entre sí, ya que aparece términos en las ecuaciones que son productos de concentraciones de portadores multiplicados por concentraciones de fotones.

6.3. Característica estática del láser

Queremos determinar en este apartado la característica estática del láser, es decir, la relación que hay, en régimen permanente, entre la potencia emitida por el láser y la corriente que lo atraviesa. Para deducirla vamos a anular las derivadas temporales en la ecuación (6.8) y definimos la constante $A := \frac{c_\circ g_n}{n}$ con lo que queda:

[6] De nuevo estamos suponiendo que la recombinación de portadores se puede describir por medio de un tiempo τ_{esp} independiente de N y no consideramos las recombinaciones no radiativas (recombinaciones en trampas y recombinaciones Auger) que pueden ser del mismo orden de magnitud que las recombinaciones radiativas.

$$
\begin{aligned}
0 &= \Gamma A P (N - N_\circ) - \frac{P}{\tau_p} + \beta \frac{\Gamma N}{\tau_{esp}} \\
0 &= \frac{I}{qV} - P_\circ A (N - N_\circ) - \frac{N}{\tau_{esp}}
\end{aligned}
\tag{6.9}
$$

En la fig. {6.4} se ha representado, junto al espectro de emisión de un láser de cavidad Fabry-Pérot guiado por el índice, su característica estática. Vemos que hay dos zonas claramente diferenciadas en las que la potencia óptica crece linealmente respecto a la corriente inyectada con pendientes muy distintas. En la primera (zona I), la pendiente es muy suave, debido a que la luz se genera, en su mayor parte, por emisión espontánea. Determinaremos la relación entre P e I en esta zona despreciando los fenómenos estimulados (ya que el medio no tiene una ganancia apreciable) y reteniendo sólo el término de emisión espontánea. En la segunda región (zona II) los fenómenos estimulados son los responsables de la generación de la mayor parte de la potencia óptica emitida. En el codo de transición entre la zona I y la II, no se puede hacer esta aproximación y, además, habría que modificar la ecuación de balance del número de fotones para predecir con exactitud el comportamiento del láser en esta zona [24]. En esta región de transición, la emisión espontánea pasa de estar repartida entre un enorme número de modos longitudinales a concentrarse en uno sólo (si el láser es monomodo) y en unos cuantos (si el láser es multimodo). Un tratamiento exacto de la emisión espontánea hace innecesaria esta linealización por tramos de la característica estática del láser. Desgraciadamente un tratamiento de esta naturaleza se sale de los límites de este libro.

6.3.1. Modelo para la zona I (emisión espontánea dominante)

El modelo para esta zona de funcionamiento lo determinamos despreciando en (6.9) los términos estimulados, con lo que resulta:

$$
\left.
\begin{aligned}
0 &= -\frac{P}{\tau_p} + \beta \Gamma \frac{N}{\tau_{esp}} \\
0 &= \frac{I}{qV} - \frac{N}{\tau_{esp}}
\end{aligned}
\right\} \Rightarrow P = \beta \frac{\Gamma \tau_g}{qV} I
\tag{6.10}
$$

Vemos que, efectivamente, la pendiente de la recta que relaciona P con I es muy pequeña, ya que la constate β es $\beta \ll 1$.

6.3.2. Modelo para la zona II (emisión espontimulada dominante)

Ahora, al sustituir en (6.9) despreciamos el término en el que aparece β con lo que nos queda:

$$\left.\begin{aligned} 0 &= \frac{\Gamma P}{\tau_g} - \frac{P}{\tau_p} \\ 0 &= \frac{I}{qV} - \frac{P}{\tau_g} - \frac{N}{\tau_{esp}} \end{aligned}\right\} \Rightarrow \left.\begin{aligned} \tau_g &= \frac{n}{c_\circ g_n (N - N_\circ)} = \Gamma \tau_p \\ N &= N_{th} := N_\circ + \frac{n}{\Gamma c_\circ g_n} \end{aligned}\right\} \Rightarrow \left.\begin{aligned} P &= \frac{\Gamma \tau_p}{qV}(I - I_{th}) \\ I_{th} &:= \frac{N_{th}}{\tau} \end{aligned}\right\} \quad (6.11)$$

Vemos que en todos los puntos de la zona II de la característica estática del láser la concentración de portadores por encima del equilibrio térmico es constante e igual a N_{th}. Cualquier incremento en la corriente inyectada, se traduce en un incremento del número de fotones que hay en la cavidad y, por tanto, en un incremento de la potencia óptica que genera el láser. Vemos, también, que en todos los puntos de la zona II la ganancia del medio amplificador permanece constante e igual a las pérdidas que hay en la cavidad resonante [7]. Sólo en la zona I de funcionamiento, un incremento en la corriente inyectada se traduce en un aumento de la concentación de portadores en la región activa. Por consiguiente, sólo en la zona I se consigue que la concentración de portadores aumente y que el coeficiente de atenuación disminuya a medida que aumenta la corriente y que, eventualmente, se haga positivo y se traduzca en una coeficiente de ganancia por unidad de longitud. Cuando la corriente llega al valor de la corriente umbral, estamos justo en el codo que une las zonas I y II. A partir de ese instante, aunque aumentemos la corriente, la concentración de portadores se mantiene igual a N_{th}. También vemos que la pendiente que relaciona P con I en la zona II es mucho mayor que en la zona I, ya que β ha desaparecido del coeficiente de poporcionalidad.

6.4. Característica de pequeña señal del láser

En este apartado vamos a determinar el modelo de pequeña señal de un láser polarizado en un punto de la zona II de funcionamiento. Vamos a derivar este modelo, no porque nos interese el resultado en sí mismo, sino porque nos va a dar una información muy útil acerca del comportamiento del láser en conmutación. Conmutamos un láser semiconductor cuando variamos rápidamente su punto de polarización desde uno en el que genera una potencia muy baja (se transmite un "0" lógico) a otro donde la potencia generada es alta (y en el que se transmite un "1" lógico). Este es la manera más sencilla de transmitir información digital con un láser y se conoce con el nombre de modulación directa [8]. El circuito con el que se polarizaría el láser para medir su característica de pequeña señal se muestra en la fig. {6.8}. En las ecuaciones que siguen vamos a despreciar la potencia óptica generada por emisión espontánea. Para el lector interesado en un modelo más detallado del láser semiconductor, lo referimos a [4]–[24]–[74].

La función de transferencia que deseamos obtener (que caracteriza el comportamiento del láser de pequeña señal) es la que relaciona $\tilde{p}$ con $\tilde{\imath}$, es decir, las fluctuaciones senoidales de la potencia óptica, que es la variable controlada, con respecto a las fluctuaciones de la corriente, que es la variable de control. La potencia $P_\circ$, la concentración N_{th} y la corriente

[7]Lo mismo pasa en los láseres de niveles discretos y, en general, en cualquier tipo de láser.

[8]Las tasas binarias máximas que se pueden transmitir con modulación directa son del orden de un par de gigabits por segundo.

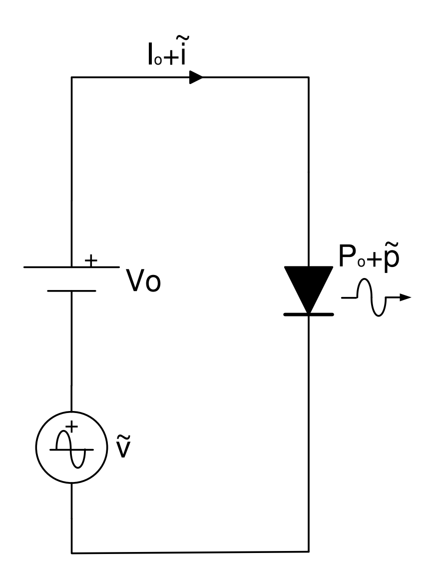


Figura 6.8: Circuito de pequeña señal del láser.

$I_\circ$ definen el punto de polarización de la zona II alrededor del cual se van a producir las perturbaciones senoidales. Por tanto verifican las siguientes igualdades:

$$\left.\begin{aligned} 0 &= \frac{P_\circ \Gamma}{\tau_g} - \frac{P_\circ}{\tau_p} \\ 0 &= \frac{I_\circ}{qV} - \frac{N_{th}}{\tau_{esp}} - \frac{P_\circ}{\tau_g} \end{aligned}\right\} \tag{6.12}$$

Ahora tenemos que sustituir en las ecuaciones dinámicas del láser los valores $P = P_\circ + \tilde{p}$, $I = I_\circ + \tilde{i}$ y $N = N_{th} + \tilde{n}$, donde $\tilde{p}$, $\tilde{i}$ y $\tilde{n}$ son *pequeñas perturbaciones senoidales* (o fasores de pequeña amplitud, si se trabaja con números complejos, cosa que haremos):

$$\left.\begin{aligned} \frac{dN}{dt} &= \frac{I}{qV} - A(N - N_\circ)P - \frac{N}{\tau_{esp}} \\ \frac{dP}{dt} &= \Gamma A(N - N_\circ)P - \frac{P}{\tau_p} \end{aligned}\right\} \Rightarrow$$

$$\left.\begin{aligned} j\omega\tilde{n} &= \frac{I_\circ}{qV} - A(N_{th} - N_\circ)P_\circ - \frac{N_{th}}{\tau_{esp}} + \frac{\tilde{i}}{qV} - A(N_{th} - N_\circ)\tilde{p} - AP_\circ\tilde{n} - A\tilde{n}\tilde{p} - \frac{\tilde{n}}{\tau_{esp}} \\ j\omega\tilde{p} &= \Gamma A(N_{th} - N_\circ)P_\circ - \frac{P_\circ}{\tau_p} + \tilde{p}[\Gamma A(N_{th} - N_\circ) - \frac{1}{\tau_p}] + \Gamma AP_\circ\tilde{n} \end{aligned}\right\} \tag{6.13}$$

Si ahora tenemos en cuenta las relaciones (6.12), que $\frac{1}{\tau_p} = \Gamma A(N_{th} - N_\circ)$ y sustituimos

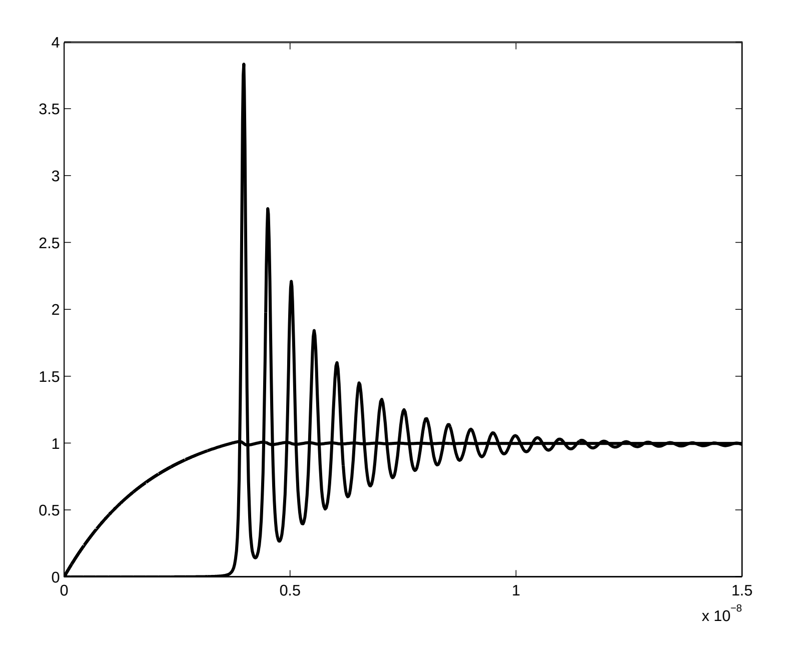


Figura 6.9: Comportamiento del láser en pequeña señal. La concentración de portadores se ha normalizado con respecto a $1,7 \times 10^{24} m^{-3}$ y la potencia con respecto a $0,65mW$. Los datos del láser se han tomado del Problema 9. Simulación realizada por Alejandro Carballar Rincón.

en (6.13), nos queda:

$$\left.\begin{aligned} j\omega\tilde{n} &= \frac{\tilde{i}}{qV} - \frac{\tilde{p}}{\tau_p} - AP_\circ\tilde{n} - A\tilde{n}\tilde{p} - \frac{\tilde{n}}{\tau_{esp}} \\ j\omega\tilde{p} &= \tilde{p}[A(N_{th} - N_\circ) - \frac{1}{\tau_p}] + AP_\circ\tilde{n} \end{aligned}\right\} \qquad (6.14)$$

Las relaciones (6.14) se pueden simplificar aún más si tenemos de nuevo en cuenta que $\Gamma A(N_{th} - N_\circ) = \frac{1}{\tau_p}$ y que dado que los fasores $\tilde{n}$ y $\tilde{p}$ son de muy pequeña amplitud, el producto de los dos se puede despreciar, con lo que queda:

$$\left.\begin{aligned} j\omega\tilde{n} &= \frac{\tilde{i}}{qV} - \frac{\tilde{p}}{\tau_p} - AP_\circ\tilde{n} - \frac{\tilde{n}}{\tau_{esp}} \\ j\omega\tilde{p} &= \Gamma AP_\circ\tilde{n} \end{aligned}\right\} \qquad (6.15)$$

Si ahora eliminamos en la ecuación (6.15) la variable $\tilde{n}$ y definimos la pulsación característica como $\omega_r := \sqrt{\frac{\Gamma AP_\circ}{\tau_p}}$ y el tiempo de amortiguamiento como $\tau_r := \frac{1}{AP_\circ + \frac{1}{\tau_{esp}}}$ deducimos que la función de transferencia del modelo de pequeña señal del láser viene dada por la siguiente expresión:

$$\left.\begin{array}{c} \dfrac{\tilde{p}}{\tilde{i}} = \dfrac{\tau_p}{qV}\dfrac{\omega_r^2}{\omega_r^2 - \omega^2 + 2j\frac{\omega}{\tau_r}} \\[2ex] \omega_r := \sqrt{\dfrac{\Gamma A P_\circ}{\tau_p}} \quad \tau_r := \dfrac{1}{AP_\circ + \frac{1}{\tau_{esp}}} \end{array}\right\} \tag{6.16}$$

Lo interesante de este resultado es que la frecuencia de oscilación resultante y el tiempo de establecimiento (o amortiguamiento, como se prefiera llamar) coinciden con los medidos en un láser en conmutación, tal y como se observa en la fig. {6.9}, donde se representa la evolución de los términos normalizados N(t) (concentración de portadores en la región activa) y P(t) (potencia óptica generada por el láser) al conmutar bruscamente la corriente desde el valor I=0 hasta un valor I=10.5mA (el resto de datos están tomados del problema resuelto 9), situado dentro de la zona II.

6.5. Problemas

Problema 8: Se tiene un láser de cavidad Fabry-Pérot cuya región activa tiene las siguientes dimensiones: L=300μm, W=1,2μm y e=0,2μm. Los coeficientes de reflexión en potencia son $R_1 = R_2 = 0,3$, el coeficiente de pérdidas por scattering Rayleigh es $\alpha_r = 40cm^{-1}$ y los índices de refracción de fase y grupo son $n_\phi = 3,5$ y $n_g = 4$. La ganancia diferencial debida a los fenómenos estimulados es $\Gamma g_n = 1,5 \times 10^{-16} cm^2$ y la densidad de portadores en la región activa bajo condiciones de transparencia es $N_\circ = 10^{18} cm^{-3}$. La tasa de emisión espontánea es $\beta = 10^{-4}$ y el tiempo medio de vida de los portadores en la región activa es $\tau = 2ns$.
Se pide:
1.- Calcular el tiempo medio de permanencia de un fotón en la cavidad resonante.
2.- Determinar la concentración de portadores y la corriente en la zona umbral.
3.- Determinar la concentración de fotones en la cavidad resonante, si se polariza el láser con una corriente de 13,834mA
4.- Deterinar la frecuencia de oscilación del láser y su tiempo de establecimiento.
5.- Estimar la tasa binaria máxima que se puede transmitir con el láser mediante modulación directa.

Solución:

1. Tal como hemos visto en el tema anterior, el tiempo medio de permanencia de un fotón en la cavidad es $\tau_p = \frac{n_g}{c_\circ(\alpha_E + \alpha_{sc})}$. $\alpha_{sc} = 40cm^{-1} = 4000m^{-1}$. Por otra parte sabemos que $\alpha_E = \frac{1}{L} Ln(\frac{1}{R_1 R_2}) = 4013m^{-1}$, con lo que las pérdidas totales por unidad de longitud valen: $\alpha = \alpha_E + \alpha_{sc} = 8013m^{-1}$. Por tanto el tiempo medio de permanenica de un fotón en la cavidad fría es $\tau_p = \frac{n_g}{c_\circ(\alpha_E + \alpha_{sc})} = \frac{4}{3 \times 10^8 \times 8013} = 1,664ps$.

2. La concentración de portadores en la zona umbral del láser la calculamos imponiendo que el coeficiente de ganancia por unidad de longitud (factor de confinamiento incluido), sea igual a las pérdidas por unidad de longitud de la cavidad fría, es decir: $g_n(N_{th} - N_\circ) = \alpha_E + \alpha_{sc}$. Sustituyendo, queda: $1,5 \times 10^{-16}(n - 10^{18}) = 80,13 \Rightarrow N_{th} = \frac{80,13}{1,5 \times 10^{-16}} + 10^{18} = 1,5342 \times 10^{18} cm^{-3}$. Para calcular la corriente umbral, tenemos en cuenta que $I_{th} = \frac{qN_{th}V}{\tau} = \frac{1,6 \times 10^{-19} 1,5342 \times 10^{24} 300 \times 10^{-6} 1,2 \times 10^{-6} 0,2 \times 10^{-6}}{2 \times 10^{-9}} = 8,834mA$.

3. La concentración de fotones dentro de la cavidad viene dada por la expresión $P = \frac{\tau_p}{qV}(I - I_{th})$. Como ya hemos calculado todos los términos que intervienen en esta expresión, podemos sustituirlos: $P = \frac{1,664 \times 10^{-12} 5 \times 10^{-3}}{1,6 \times 10^{-19} 300 \times 10^{-6} 1,2 \times 10^{-6} 0,2 \times 10^{-6}} = 7,22 \times 10^{20} m^{-3}$.

4. Sabemos que $\omega_r^2 = \frac{AP_\circ}{\tau_p} = \frac{c_\circ g_N P_\circ}{n_g \tau_p} = \frac{3 \times 10^8 1,5 \times 10^{-20} 7,22 \times 10^{20}}{41,664 \times 10^{-12}} \Rightarrow f_r = 3,5163GHz$. $\frac{2}{\tau_r} = \frac{c_\circ g_N P_\circ}{n_g} + \frac{1}{\tau} = \frac{3 \times 10^8 1,5 \times 10^{-20} 7,22 \times 10^{20}}{4} + \frac{1}{2 \times 10^{-9}} \Rightarrow \tau_r = 1,524ns$.

5. Vamos a admitirque el pulso se hace estable cuando transcurre un tiempo que es dos veces el de establecimiento. Por tanto, si asumimos una codificación NRZ, tendremos que $B_{max} \simeq \frac{1}{21,524 \times 10^{-9}} = 328,1MHz$.

Capítulo 7

El láser de cavidad distribuida: modos acoplados

Resumen: En este capítulo vamos a estudiar el láser de cavidad distribuida utilizando el formalismo de modos acoplados. Sin embargo la deducción que vamos a hacer de las ecuaciones no es matemáticamente rigurosa. Esto se hace así porque la deducción rigurosa y correcta de las ecuaciones de modos acoplados es larga y requiere unos conocimientos matemáticos que no se le presuponen al lector. La manera en que vamos a razonar es "intuitiva" y nos parece que esta razón justifica la falta de rigor matemático. En primer lugar explicaremos, de una manera también cualitativa, por qué una perturbación senoidal del índice de refracción puede acoplar energía entre dos modos propios que se propagan en sentidos contrarios. A continuación "deduciremos" la ecuación de modos acoplados para, a continucación, resolverla y demostrar que el láser propuesto es, en principio, bimodal, es decir, que emite a dos longitudes de onda distintas. Veremos, por último, qué modificaciones hay que introducir en la estructura para que el dispositivo emita a una única longitud de onda.

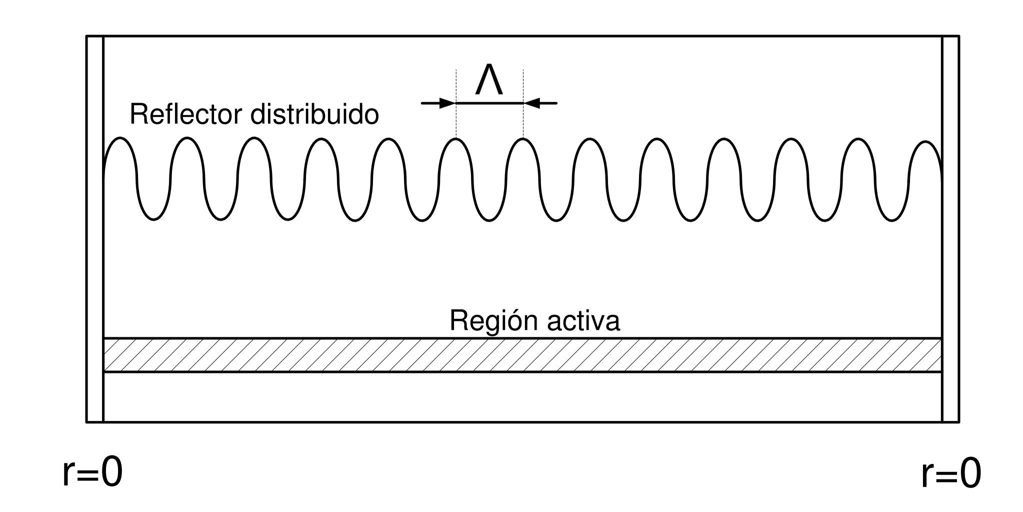


Figura 7.1: Corte longitudinal de un láser de Bragg.

7.1. Introducción

En la fig. {7.1} se muestra una sección longitudinal de un láser de cavidad distribuida. Las superficies anterior y posterior del dispositivo han sido tratadas de manera que tengan un coeficiente de reflexión en campo nulo. Sin embargo, para longitudes de onda cercanas a un determinado valor (longitud de onda de Bragg) la perturbación periódica senoidal presente en el dispositivo introduce un fuerte acoplamiento (intercambio de energía) entre dos modos que se propagan en sentidos opuestos [56].

En el láser el índice de refracción complejo viene descrito por medio de la siguiente expresión:

$$n(z) = n_\circ + n_1 cos(\frac{2\pi}{\Lambda} z) + jn_2 \tag{7.1}$$

donde $n_\circ$ es el índice de refracción promedio, n_1 es la amplitud de la perturbación del índice, Λ es el periodo de la perturbación y n_2 representa la ganancia del medio (de la que es responsable, por supuesto, la región activa del láser). Supongamos que por el dispositivo se está propagando una onda y que en los máximos del índice de refracción se producen reflexiones parciales, tal como se muestra en la fig. {7.2}.

La longitud de onda para la cual la estructura presenta un coeficiente de reflexión elevado es aquélla para la que todas las ondas parcialmente reflejada se suman en fase, es decir, para la que se cumple:

$$\Delta\Phi = \frac{2\Lambda}{\lambda_B/n_\circ} 2\pi = 2\pi \Rightarrow \lambda_B = 2n_\circ\Lambda \tag{7.2}$$

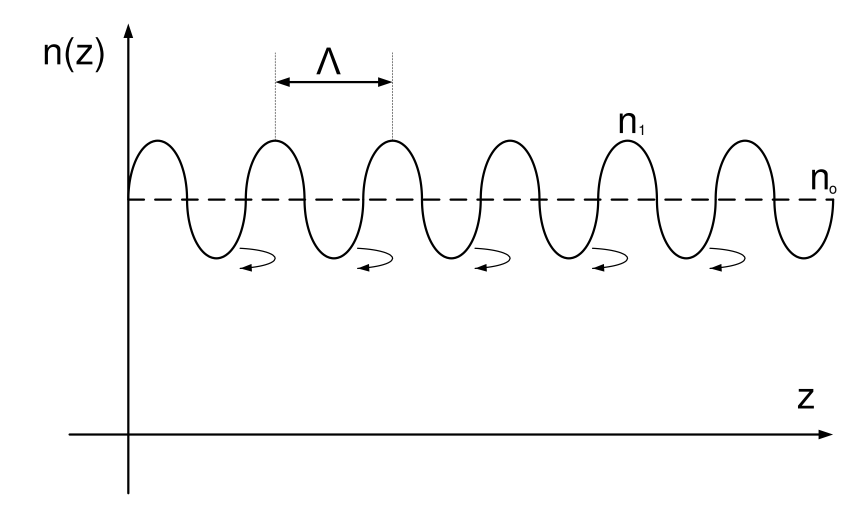


Figura 7.2: Refexiones parciales en una red de difracción de Bragg.

donde no tiene sentido que nos planteemos que el desfase sea un múltiplo entero de 2π distinto de la unidad, ya que a esas longitudes de onda el medio activo no sería capaz de amplificar la luz. Por tanto, si suponemos que las ondas que se propagan son planas (en realidad lo único que hacemos es ignorar la distribución transversal del modo guiado) llegamos a la conclusión de que estas ondas tienen constantes de propagación cercanas a $\pm n_\circ \frac{2\pi}{\lambda_B} = \pm\frac{\pi}{\Lambda}$. Por consiguiente, y dado que la red de difracción de Bragg tiene una longitud finita, lo único que podemos afirmar de una manera cierta es que por el láser sólo se puede emitir ondas electromagnéticas cuyas constantes de propagación en el medio tomen valores cercanos a $\pm\frac{\pi}{\Lambda}$:

$$\beta_\lambda \simeq \beta_B = \pm\frac{\pi}{\Lambda} \tag{7.3}$$

La longitud de onda real de emisión del láser, aunque próxima a λ_B, no tiene por qué coincidir exactamente con este valor. Vamos a definir una serie de constantes de propagación asociadas a los distintos índices que aparecen en la estructura, que nos resultarán muy útiles más adelante [34]:

$$\begin{cases} \delta := \beta_B - \beta_\lambda \\ \kappa := \dfrac{k_\circ n_1}{2} \\ g := 2k_\circ n_2 \end{cases} \tag{7.4}$$

El término δ representa la desviación, medida como una constante de propagación, entre la longitud de onda de la luz que se propaga por el medio y la longitud de onda de Bragg. La constante κ es la amplitud de la perturbación del índice (n_1) expresada como una

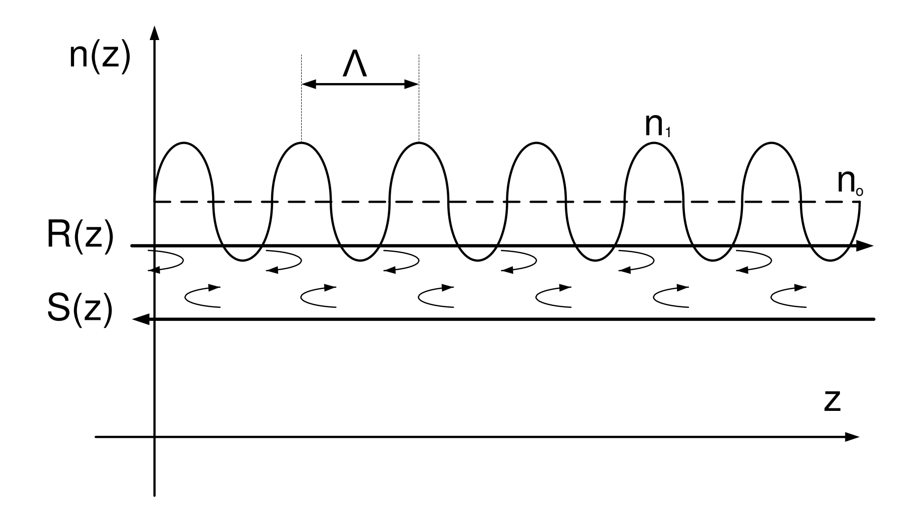


Figura 7.3: Intercambio de enrgía entre las ondas contrapropagadas en el reflector de Bragg.

constante de propagación e indica cómo de intenso es el acoplamiento entre las ondas co y contrapropagadas. Por último, g es la ganancia del medio medida, también, como constante de propagación.

7.2. Ecuación de Helmholtz

Vamos a plantear ahora las ecuaciones de dos ondas planas que se propagan por un medio como el descrito por la fig. {7.3}. Evidentemente, la perturbación senoidal del índice provoca un intercambio de energía entre dos ondas que se propagan en sentidos opuestos con una misma constante de propagación de fase (en valor absoluto) cercana a β_B. La amplitud de la onda que viaja hacia la derecha la denotaremos $R(z)$ y asumimos (dado que la perturbación del índice es pequeña) que varía lentamente frente a la coordenada longitudinal z. La amplitud de la onda que se propaga hacia la izquierda es $S(z)$ y cambia, por idénticos motivos, lentamente con respecto a la coordenada longitudinal z.

La expresión analítica del índice complejo de refracción de la estructura es:

$$n(z) = n_\circ + n_1 \frac{e^{j\frac{2\pi n}{\Lambda}} + e^{-j\frac{2\pi n}{\Lambda}}}{2} + jn_2 \tag{7.5}$$

y es precisamente el término en n_1 el que acopla las ondas $R(z)$ y $S(z)$. La ecuación que describe la propagación de las ondas es la ecuación de Helmholtz:

$$\frac{d^2E}{dz^2} + k_\circ^2 n(z)E = 0 \tag{7.6}$$

Postulamos como solución de esta ecuación la suma de dos ondas que tienen la misma

constante de propagación, β_λ en valor absoluto, pero que se propagan en sentidos contrarios. Además vamos a considerar que las envolventes $R(z)$ y $S(z)$ varían lentamente con la coordenada longitudinal, es decir, ensayamos con una solución cuya expresión es:

$$\begin{aligned} E(z) &= R(z)e^{-j\beta_\lambda z} + S(z)e^{j\beta_\lambda z} \Rightarrow \\ \dot{E}(z) &= (\dot{R} - j\beta_\lambda R)e^{-j\beta_\lambda z} + (\dot{S} + j\beta_\lambda S)e^{j\beta_\lambda z} \Rightarrow \\ \ddot{E}(z) &\simeq -(2j\beta_\lambda \dot{R} + \beta_\lambda^2 R)e^{-j\beta_\lambda z} + (2j\beta_\lambda \dot{S} - \beta_\lambda^2 S)e^{j\beta_\lambda z} \end{aligned} \tag{7.7}$$

En esta última expresión hemos tenido en cuenta que las envolventes varían muy poco a lo largo de una longitud de onda, por lo que los términos cuadráticos y lineales en β_λ son dominantes frente al resto.

Ahora vamos a calcular $k_\circ^2 n^2$ despreciando los términos que sean productos de dos magnitudes pequeñas (tanto n_1 como n_2 son pequeñas perturbaciones con respecto al valor medio del índice de refracción $n_\circ$):

$$\begin{aligned} & k_\circ^2 n^2(z) \simeq k_\circ^2 \left(n_\circ^2 + 2jn_\circ n_2 + 2n_\circ n_1 \frac{e^{2j\beta_B z} + e^{-2j\beta_B z}}{2} \right) \\ & E(z) = R(z)e^{-j\beta_\lambda z} + S(z)e^{j\beta_\lambda z} \Rightarrow \\ & k_\circ^2 n^2(z)E(z) \simeq \beta_B^2 Re^{-j\beta_\lambda z} + \beta_B^2 Se^{j\beta_\lambda z} + 2j\beta_B k_\circ n_2 Re^{-j\beta_\lambda z} + 2j\beta_B k_\circ n_2 Se^{j\beta_\lambda z} + \\ & + \beta_B k_\circ n_1 Re^{j(2\beta_B - \beta_\lambda)z} + \beta_B k_\circ n_1 Se^{-j(2\beta_B - \beta_\lambda)z} \simeq \beta_B^2 Re^{-j\beta_\lambda z} + \beta_B^2 Se^{j\beta_\lambda z} + \\ & + 2j\beta_B k_\circ n_2 Re^{-j\beta_\lambda z} + 2j\beta_B k_\circ n_2 Se^{j\beta_\lambda z} + \beta_B k_\circ n_1 Re^{j\beta_\lambda z} + \beta_B k_\circ n_1 Se^{-j\beta_\lambda z} \end{aligned} \tag{7.8}$$

Los términos en $e^{j(2\beta_B+\beta_\lambda)z}$ y $e^{-j(2\beta_B+\beta_\lambda)z}$ no se han considerado ya que para las longitudes de onda correspondientes no hay realimentación alguna y el medio activo es absorbente. Además hemos hecho la aproximación $2\beta_B - \beta_\lambda \simeq \beta_\lambda$. La expresión resultante es la suma de dos ondas, una que se propaga de izquierda a derecha y otra que se propaga de derecha a izquierda. Para que la suma de ambas ondas sea nula, el sumatorio de los coeficientes de las ondas que se propagan de izquierda a derecha ha ser igual a cero al igual que el sumatorio de los coeficientes de las ondas que se propagan en sentido contrario. Por consiguiente, tienen que verificarse las siguientes relaciones:

$$\begin{aligned} (\beta_B^2 - \beta_\lambda^2)R + 2j\beta_B k_\circ n_2 R - 2j\beta_\lambda \dot{R} + \beta_B k_\circ n_1 S = 0 \\ (\beta_B^2 - \beta_\lambda^2)S + 2j\beta_B k_\circ n_2 S - 2j\beta_\lambda \dot{S} + \beta_B k_\circ n_1 R = 0 \end{aligned} \tag{7.9}$$

Si ahora dividimos ambas ecuaciones por $2j\beta_\lambda$, aparecen los siguientes coeficientes:

$$\frac{\beta_B^2-\beta_\lambda^2}{2j\beta_\lambda}=\frac{(\beta_B+\beta_\lambda)(\beta_B-\beta_\lambda)}{2j\beta_\lambda}\simeq\frac{2\beta_\lambda(\beta_B-\beta_\lambda)}{2j\beta_\lambda}=-j\delta$$

$$\frac{2j\beta_B-k_\circ n_2}{2j\beta_\lambda}\simeq k_\circ n_2=\frac{g}{2} \tag{7.10}$$

$$\frac{\beta_B-k_\circ n_1}{2j\beta_\lambda}\simeq -j\frac{k_\circ n_1}{2}=-j\kappa$$

El sistema de ecuaciones diferenciales resultante es:

$$-\dot{R}+\left(\frac{g}{2}-j\delta\right)R=j\kappa S$$

$$\dot{S}+\left(\frac{g}{2}-j\delta\right)S=j\kappa R \tag{7.11}$$

donde g y κ son la constante de ganancia y acoplo, mientras que δ es la variable que representa la longitud de onda o, si se prefiere, la constante de propagación. Estos términos se definieron en (7.4).

Las ecuaciones (7.11) definen un sistema de dos ecuaciones diferenciales de primer orden acopladas entre sí para el que hay procedimientos de integración directa. En vez de aplicarlos, vamos a deducir dos ecuaciones diferenciales de segundo orden, lineales y *desacopladas*. Evidentemente, en la solución de estas ecuaciones aparecen cuatro constantes de integración de las que sólo dos son independientes. Este problema lo resolveremos imponiendo que las soluciones de estas ecuaciones diferenciales de segundo grado sean también soluciones de (7.11).

$$\left.\begin{array}{r}-\dot{R}+\left(\frac{g}{2}-j\delta\right)R=j\kappa S\\ \dot{S}+\left(\frac{g}{2}-j\delta\right)S=j\kappa R\end{array}\right\}\Rightarrow\left.\begin{array}{r}-\ddot{R}+\left(\frac{g}{2}-j\delta\right)\dot{R}=j\kappa\dot{S}\\ \ddot{S}+\left(\frac{g}{2}-j\delta\right)\dot{S}=j\kappa\dot{R}\end{array}\right\}\Rightarrow$$

$$\left.\begin{array}{r}-\ddot{R}+\left(\frac{g}{2}-j\delta\right)\dot{R}=-\kappa^2R-j\kappa\left(\frac{g}{2}-j\delta\right)S\\ \ddot{S}+\left(\frac{g}{2}-j\delta\right)\dot{S}=\kappa^2R+j\kappa\left(\frac{g}{2}-j\delta\right)R\end{array}\right\}\Rightarrow\left.\begin{array}{r}\ddot{R}-\gamma^2R=0\\ \ddot{S}-\gamma^2S=0\end{array}\right\} \tag{7.12}$$

Donde la constante γ verifica la relación $\gamma^2:=\kappa^2+(\frac{g}{2}-j\delta)^2$. Ahora podemos integrar de manera inmediata las ecuaciones diferenciales (7.12), cuya solución es:

$$R(z)=r_1e^{\gamma z}+r_2e^{-\gamma z}$$

$$S(z)=s_1e^{\gamma z}+s_2e^{-\gamma z} \tag{7.13}$$

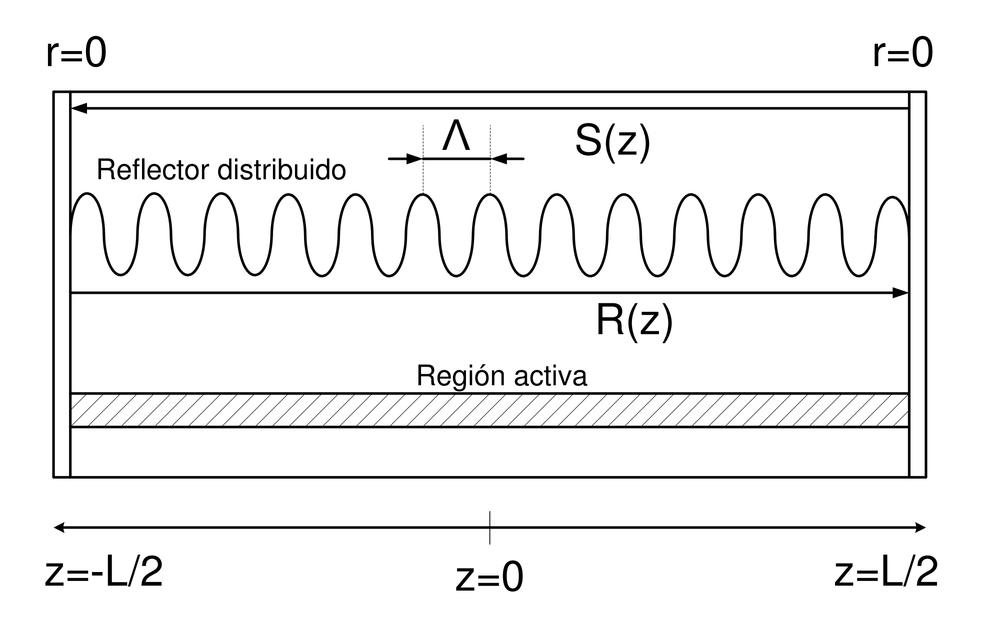


Figura 7.4: Sistema de referencia para imponer las conciciones de contorno r=0.

Si ahora imponemos que las soluciones (7.13) verifiquen las ecuaciones (7.11) nos quedan las siguientes relaciones entre las constantes de integración r_1, r_2, s_1 y s_2:

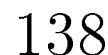
$$\begin{aligned} -\gamma r_1 + \left(\frac{g}{2} - j\delta\right) r_1 &= j\kappa s_1 \\ \gamma r_2 + \left(\frac{g}{2} - j\delta\right) r_2 &= j\kappa s_2 \end{aligned} \tag{7.14}$$

La pregunta que ahora debemos plantearnos es cómo podemos hallar otras dos relaciones que nos permitan determinar los valores de las cuatro constantes de integración. Si miramos la fig. {7.1} vemos que los coeficientes de reflexión en las dos superficies extremas del láser son iguales a cero. En la fig. {7.4} vemos claramente que para que el coeficiente de reflexión sea cero en la superficie izquierda, la onda R debe anularse en $z = \frac{-L}{2}$. De igual manera, para que el coeficiente de reflexión en la superficie de la derecha sea cero, la onda S particularizada en $z = \frac{L}{2}$ debe ser igual a cero también. De esta manera deducimos las siguientes relaciones entre las constantes de integración:

$$\begin{aligned} r_1 e^{\frac{-\gamma L}{2}} + r_2 e^{\frac{\gamma L}{2}} &= 0 \\ s_1 e^{\frac{\gamma L}{2}} + s_2 e^{\frac{-\gamma L}{2}} &= 0 \end{aligned} \tag{7.15}$$

7.3. Condición de oscilación del láser

Si ahora reunimos las relaciones (7.14) y (7.15) nos queda el siguiente sistema lineal y homogéneo de cuatro ecuaciones con cuatro incógnitas:



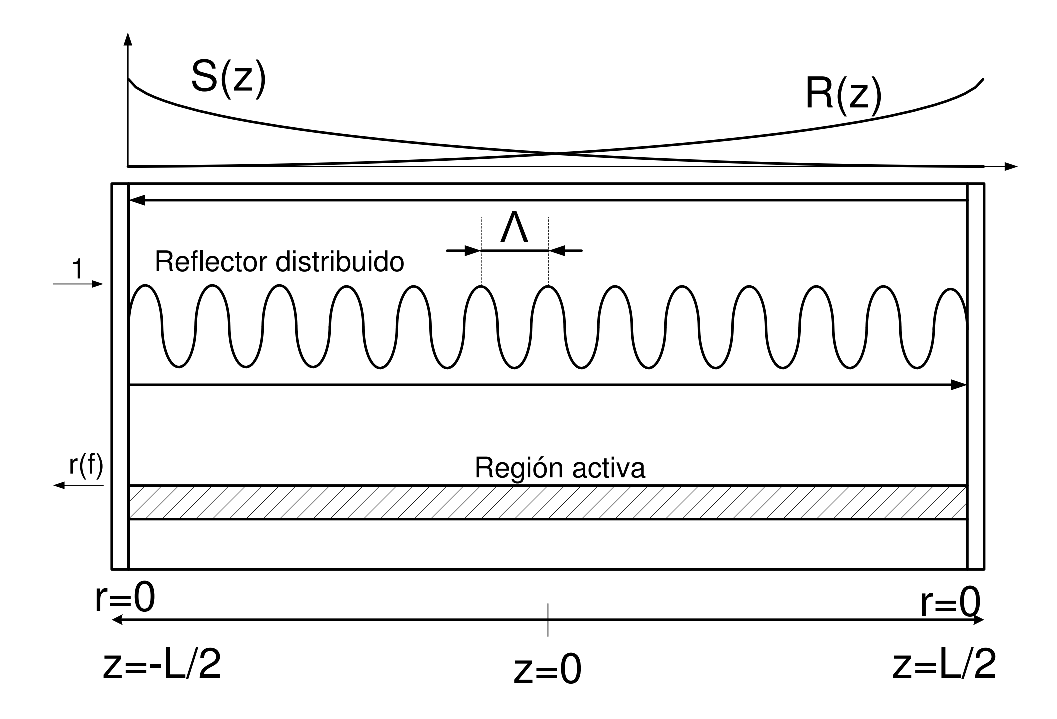


Figura 7.5: Estructura a la que vamos a calcularle el coeficiente de reflexión.

$$\begin{cases} r_1 e^{\frac{-\gamma L}{2}} + r_2 e^{\frac{\gamma L}{2}} = 0 \\ s_1 e^{\frac{\gamma L}{2}} + s_2 e^{\frac{-\gamma L}{2}} = 0 \\ -\gamma r_1 + \left(\frac{g}{2} - j\delta\right) r_1 - j\kappa s_1 = 0 \\ \gamma r_2 + \left(\frac{g}{2} - j\delta\right) r_2 - j\kappa s_2 = 0 \end{cases} \tag{7.16}$$

Este sistema siempre tiene como solución la solución trivial $r_1 = r_2 = s_1 = s_2 = 0$, es decir, el láser no emite luz alguna. Para que exista una solución distinta de la trivial, el determinante del sistema debe ser igual a cero, en cuyo caso queda un parámetro libre que fija la potencia de emisión del láser. Vamos ahora a operar sobre las filas de la matriz asociada al sistema de ecuaciones, para convertirla en una matriz triangular superior, ya para estas matrices el determinante asociado es igual al producto de los elementos de la diagonal principal:

$$\begin{bmatrix} e^{\frac{-\gamma L}{2}} & e^{\frac{\gamma L}{2}} & 0 & 0 \\ 0 & 0 & e^{\frac{\gamma L}{2}} & e^{\frac{-\gamma L}{2}} \\ -\gamma + (\frac{g}{2} - j\delta) & 0 & -j\kappa & 0 \\ 0 & \gamma + (\frac{g}{2} - j\delta) & 0 & -j\kappa \end{bmatrix} F_{2\leftrightarrow 4} \rightarrow$$

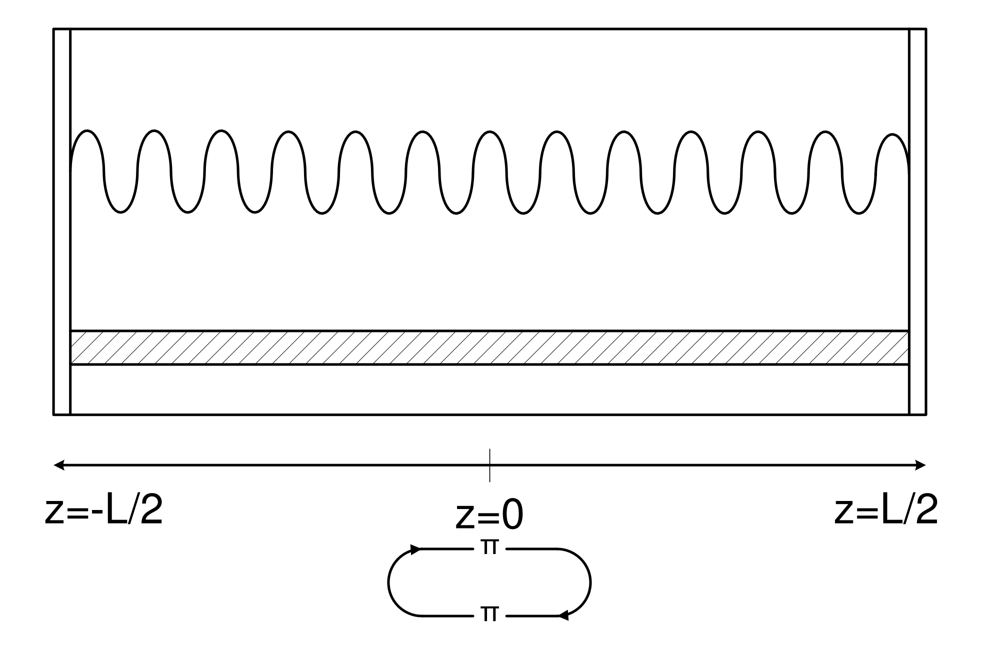


Figura 7.6: Desfases que se producen dentro de la estructura a las longitudes de onda de emisión.

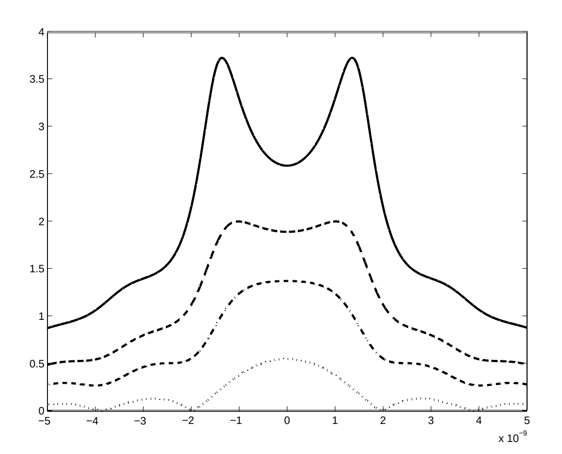
$$
\begin{bmatrix}
e^{\frac{-\gamma L}{2}} & e^{\frac{\gamma L}{2}} & 0 & 0 \\
0 & \gamma + (\frac{g}{2} - j\delta) & 0 & -j\kappa \\
-\gamma + (\frac{g}{2} - j\delta) & 0 & -j\kappa & 0 \\
0 & 0 & e^{\frac{\gamma L}{2}} & e^{\frac{-\gamma L}{2}}
\end{bmatrix}
F_{31}[\gamma - (\frac{g}{2} - j\delta)]e^{\frac{-\gamma L}{2}} \rightarrow
$$

$$
\begin{bmatrix}
e^{\frac{-\gamma L}{2}} & e^{\frac{\gamma L}{2}} & 0 & 0 \\
0 & \gamma + (\frac{g}{2} - j\delta) & 0 & -j\kappa \\
0 & e^{\gamma L}[\gamma - (\frac{g}{2} - j\delta)] & -j\kappa & 0 \\
0 & 0 & e^{\frac{\gamma L}{2}} & e^{\frac{-\gamma L}{2}}
\end{bmatrix}
F_{32}[-e^{\gamma L}\frac{\gamma - (\frac{g}{2} - j\delta)}{\gamma + (\frac{g}{2} - j\delta)}] \rightarrow
$$

$$
\begin{bmatrix}
e^{\frac{-\gamma L}{2}} & e^{\frac{\gamma L}{2}} & 0 & 0 \\
0 & \gamma + (\frac{g}{2} - j\delta) & 0 & -j\kappa \\
0 & 0 & -j\kappa & j\kappa e^{\gamma L}\frac{\gamma - (\frac{g}{2} - j\delta)}{\gamma + (\frac{g}{2} - j\delta)} \\
0 & 0 & e^{\frac{\gamma L}{2}} & e^{\frac{-\gamma L}{2}}
\end{bmatrix}
F_{43}[\frac{e^{\frac{\gamma L}{2}}}{j\kappa}] \rightarrow
$$

$$
\begin{bmatrix}
e^{\frac{-\gamma L}{2}} & e^{\frac{\gamma L}{2}} & 0 & 0 \\
0 & \gamma + (\frac{g}{2} - j\delta) & 0 & -j\kappa \\
0 & 0 & -j\kappa & j\kappa e^{\gamma L}\frac{\gamma - (\frac{g}{2} - j\delta)}{\gamma + (\frac{g}{2} - j\delta)} \\
0 & 0 & 0 & e^{\frac{\gamma L}{2}} + e^{\frac{3\gamma L}{2}}\frac{\gamma - (\frac{g}{2} - j\delta)}{\gamma + (\frac{g}{2} - j\delta)}
\end{bmatrix}
\tag{7.17}
$$

Para que el determinante de (7.17) sea nulo, se tiene que cumplir que el producto de sus elementos diagonales sea igual a cero, es decir que:



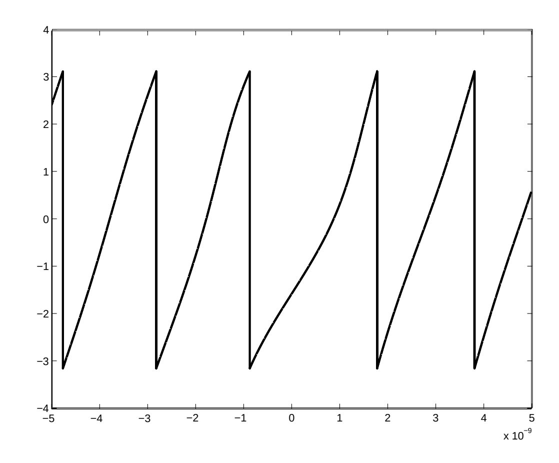


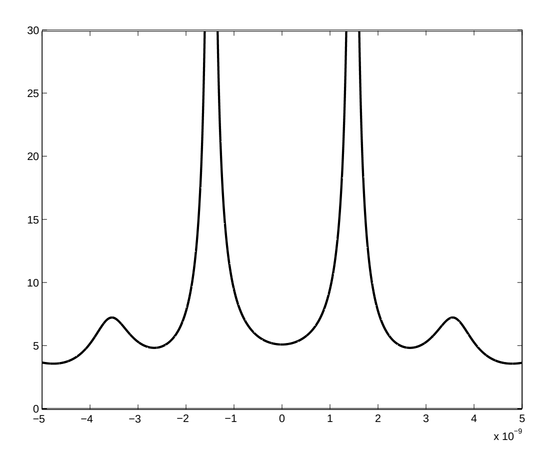
Figura 7.7: Coeficiente de reflexión complejo (módulo a la izquierda y argumento a la derecha) de una estructura láser de Bragg bimodo con ganancias inferiores a la umbral. Las curvas de módulo representadas se corresponden con las ganancias $g = 0cm^{-1}$, $g = 100cm^{-1}$,$g = 200cm^{-1}$ y $g = 250cm^{-1}$. El argumento se ha representado sólo para $g = 250cm^{-1}$. El resto de parámetros son, en todas las gráficas, los mismos: $\kappa = 50cm^{-1}$, $L = 125\mu m$ $n_\circ = 3,5$ y $\lambda_B = 1,3\mu m$. Simulación relizada por Alejandro Carballar Rincón.

$$\Delta = e^{-\frac{\gamma L}{2}}\left[\gamma + \frac{g}{2} - j\delta\right](-j\kappa)e^{\frac{\gamma L}{2}}\frac{e^{-\gamma L}\left[\gamma + \left(\frac{g}{2} - j\delta\right)\right] + e^{\frac{\gamma L}{2}}\left[\gamma - \left[\left(\frac{g}{2} - j\delta\right)\right]\right]}{\gamma + \left(\frac{g}{2} - j\delta\right)} =$$

$$= j\kappa\left[2\gamma Ch(\gamma L) - \left(\frac{g}{2} - j\delta\right)2Sh(\gamma L)\right] = 0 \Rightarrow \tag{7.18}$$

$$Cotgh(\gamma L) = \frac{\left(\frac{g}{2} - j\delta\right)L}{\gamma L}$$

La ecuación (7.18) es compleja y, de hecho, hay dos incógnitas reales: por un lado el valor de δ, es decir la longitud de onda de oscilación del láser y, por otro, $\frac{g}{2}$, es decir, el valor que debe tomar la ganacia activa para que el dispositivo pueda emitir luz. Los parámetros β_B y κ son conocidos ya que sabemos qué periodo tiene la perturbación senoidal del índice introducida en la estructura y la amplitud de la misma. Sin embargo la resolución de esta ecuación no aporta ninguna información intuitiva acerca del comportamiento del dispositivo. Por ello, vamos a considerar la fig. {7.5} y vamos a calcular el coeficiente de reflexión de la estructura representada. Esto quiere decir que vamos a resolver el mismo sistema de ecuaciones a partir del que dedujimos la condición de oscilación, salvo que ahora la condición de contorno en $z = -\frac{L}{2}$ es $S(-\frac{L}{2}) = 1$ (lanzamos, por la izquierda, una onda de amplitud 1 a todas las longitudes de onda), con lo que obtenemos el sistema:



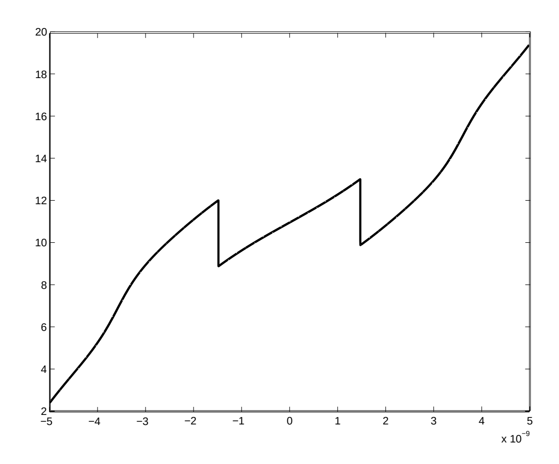


Figura 7.8: Coeficiente de reflexión complejo (módulo a la izquierda y argumento a la derecha) de una estructura láser de Bragg bimodo con ganancia igual a la umbral. La curva de módulo representadas se corresponden con la ganancia $g = 372,9cm^{-1}$. El argumento se ha representado también para $g = 372,9cm^{-1}$. En la curva del argumento se observa claramente un salto de fase igual a π a las dos longitudes de onda de emisión. El resto de parámetros son, en todas las gráficas, los mismos: $\kappa = 50cm^{-1}$, $L = 125\mu m$ $n_\circ = 3,5$ y $\lambda_B = 1,3\mu m$. Simulación realizada por Alejandro Carballar Rincón.

$$\begin{bmatrix} e^{\frac{-\gamma L}{2}} & e^{\frac{\gamma L}{2}} & 0 & 0 \\ 0 & 0 & e^{\frac{\gamma L}{2}} & e^{\frac{-\gamma L}{2}} \\ -\gamma + (\frac{g}{2} - j\delta) & 0 & -j\kappa & 0 \\ 0 & \gamma + (\frac{g}{2} - j\delta) & 0 & -j\kappa \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (7.19)$$

Para resolver este problema, realizamos las mismas operaciones por filas que cuando calculamos el determinante de la matriz, pero ahora lo hacemos con la matriz 4×5, resultante de añadir el vector columna de los términos independientes. No vamos a repetir las operaciones y damos los valores deducidos de los coeficientes s_1 y s_2 y el coeficiente de reflexión resultante:

$$\left.\begin{aligned} s_1 &= \frac{j\kappa/2e^{\frac{-\gamma L}{2}}}{\gamma Ch(\gamma L) - \left(\frac{g}{2} - j\delta\right) Sh(\gamma L)} \\ s_2 &= \frac{-j\kappa/2e^{\frac{\gamma L}{2}}}{\gamma Ch(\gamma L) - \left(\frac{g}{2} - j\delta\right) Sh(\gamma L)} \end{aligned}\right\} \Rightarrow$$

$$r = S(-L/2) = \frac{-j\kappa Sh(\gamma L)}{\gamma Ch(\gamma L) - (\frac{g}{2} - j\delta)Sh(\gamma L)} \quad (7.20)$$

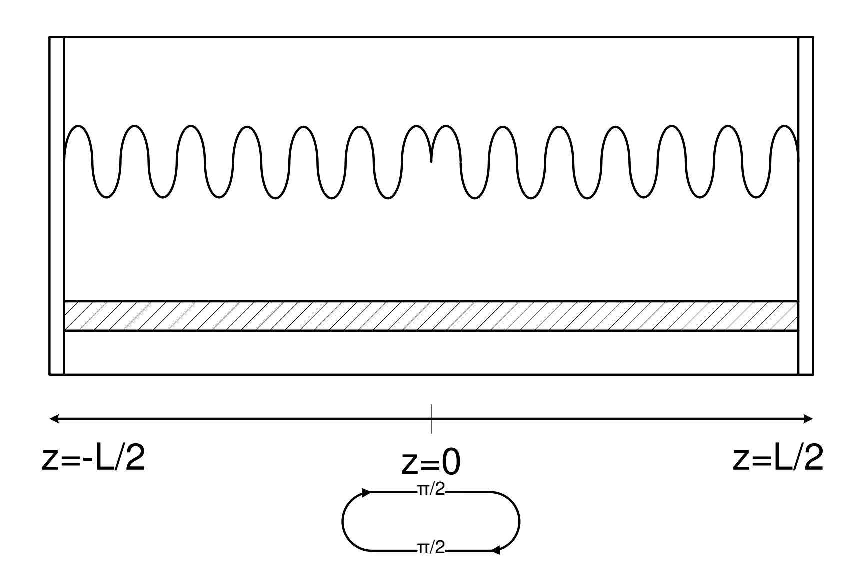
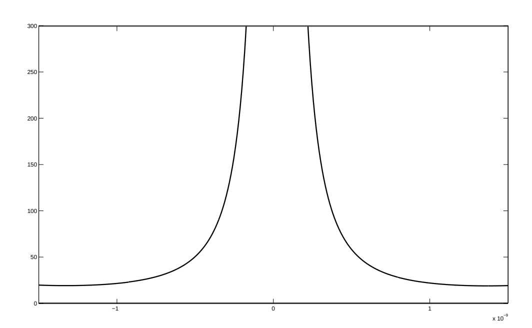


Figura 7.9: Estructura de un láser de Bragg monomodo.

Este coeficiente de reflexión es complejo, lo que indica que la onda reflejada tiene una amplitud distinta de la onda incidente y un desfase con respecto a ella. Lo que vamos a hacer es representar tanto el módulo del coeficiente de reflexión como su argumento. El único parámetro que vamos a considerar variable es λ (o, lo que es equivalente, δ). Vamos a fijar la longitud del láser igual a $L = 125\mu m$, el índice de refracción promedio lo tomamos $n_\circ = 3,5$, la longitud de onda de Bragg correspondiente a la perturbación del índice será $\lambda_B = 1,3\mu m$, el coeficiente de acoplamiento lo hacemos igua a $\kappa = 50cm^{-1}$ y parametrizamos las curvas del módulo para los valores de ganancia $g = 0cm^{-1}, 150cm^{-1}, 200cm^{-1}\ y\ 250cm^{-1}$. La gráfica asociada al argumento del coeficiente de reflexión sólo se ha representado para este último caso, tal como se muestra en la fig. {7.7}.

En estas gráficas vemos que si la ganancia del medio se hace suficientemente grande, el módulo del coeficiente de reflexión tiende a infinito a dos longitudes de onda distintas [74]. ¿Cómo debemos interpretar este hecho? Si vemos la gráfica correspondiente al argumento del coeficiente de reflexión con $g = 250cm^{-1}$, comprobamos que a la dos longitudes de onda de emisión donde el coeficiente de reflexión es máximo se produce un desfase de valor cercano a π. Si ahora simulamos el comportamiento de la estructura para $g = 372,9cm^{-1}$ vemos en la fig. {7.8} que a las dos longitudes de onda donde r se hace infinito, el desfase acumulado en una ida y vuelta a lo largo de la cavidad es exactamente igual a 2π con lo que se dan las condiciones de realimentación positiva. Sin embargo esta situación no nos interesa en absoluto, ya que lo que se quiere es que el dispositivo emita a una única longitud de onda, es decir, se busca un láser monomodo y no bimodal.

Para solventar el problema se puede modificar la estructura de la perturbación introduciendo un salto de fase de valor π en z=0. Se puede comprobar (empleando las técnicas de simulación numérica descritas en el siguiente capítulo) que el láser representado en la fig. {7.9} emite a una única longitud de onda, tal y como se muestra en la fig. {7.10}. Normalmente el salto de fase que se introduce dentro del reflector de Bragg vale $\frac{\pi}{2}$ (también conocido como salto de fase de $\frac{\lambda}{4}$) [34]–[74] ya que se consigue una pureza espectral mayor.



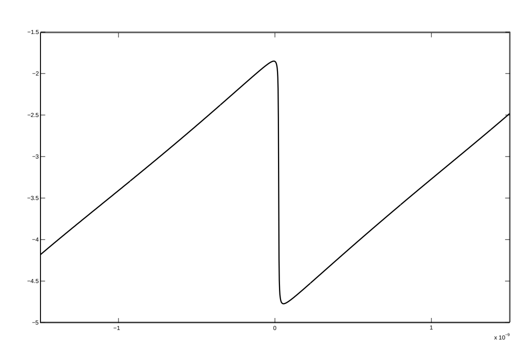


Figura 7.10: Coeficiente de reflexión complejo (módulo a la izquierda y argumento a la derecha) de una estructura láser de Bragg monomodo (con un salto de fase igual a π)) con ganancia igual a la umbral. La curva de módulo representadas se corresponden con la ganancia $g = 372,9cm^{-1}$. El argumento se ha representado también para $g = 372,9cm^{-1}$. En la curva del argumento se observa claramente un salto de fase igual a π a la longitud de onda de emisión. El resto de parámetros son, en todas las gráficas, los mismos: $\kappa = 50cm^{-1}$, $L = 125\mu m$ $n_\circ = 3,5$ y $\lambda_B = 1,3\mu m$. Simulación realizada por Alejandro Carballar Rincón.

Estamos utilizando un modelo demasiado simplificado del láser de cavidad distribuida, por lo que no podemos justificar esta afirmación.

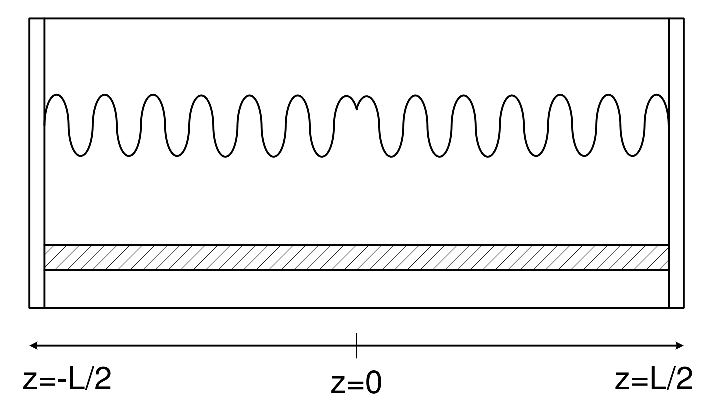
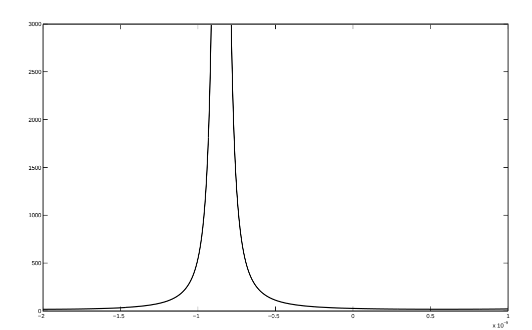


Figura 7.11: Estructura de un láser de Bragg monomodo con salto de fase de $\frac{\pi}{2}$.



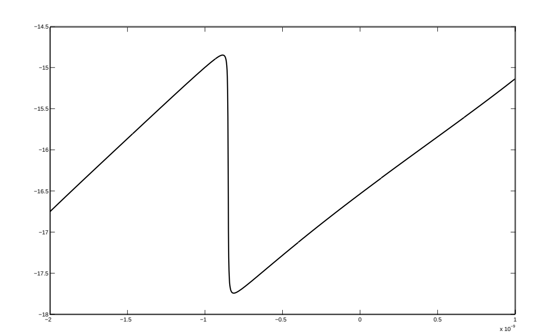


Figura 7.12: Coeficiente de reflexión complejo (módulo a la izquierda y argumento a la derecha) de una estructura láser de Bragg monomodo (con un salto de fase igual a $\frac{\pi}{2}$)) con ganancia igual a la umbral. La curva de módulo representadas se corresponden con la ganancia $g = 372,9cm^{-1}$. El argumento se ha representado también para $g = 372,9cm^{-1}$. En la curva del argumento se observa claramente un salto de fase igual a π a la longitud de onda de emisión. El resto de parámetros son, en todas las gráficas, los mismos: $\kappa = 50cm^{-1}$, $L = 125\mu m$ $n_\circ = 3,5$ y $\lambda_B = 1,3\mu m$. Simulación realizada por Alejandro Carballar Rincón.

7.4. Problemas

Problema 9: Simular el comportamiento de un láser de cavidad distribuida con un salto de fase de $\frac{\pi}{2}$. Las condiciones de simulaciones son las mismas que las del láser bimodo y el láser monomodo con un salto de fase de π. Las simulaciones deben hacerse utilizando el modelo matricial explicado en el siguiente capítulo.

Solución: Para resolver este problema modificamos la perturbación senoidal del índice con un salto de fase de valor $\frac{\pi}{2}$ en z=0. El láser resultante se ha representado en la fig. {7.11}. Para simular su comportamiento, hacemos unos cálculos numéricos descritos en el siguiente capítulo que se basan en el formalismo de las matrices de transmisión. El dispositivo emite a una única longitud de onda, tal y como se muestra en la fig. {7.12}. Se observa un desplazamiento de la longitud de emisión con respecto a la longitud de onda de Bragg.

Capítulo 8

Estudio matricial de las cavidades distribuidas

Resumen: En este capítulo vamos a estudiar de nuevo los láseres de cavidad distribuida. Esta vez utilizaremos el formalismo de las matrices de scattering y transmisión para deducir las propiedades de las cavidades resonantes de estos dispositivos. Este formalismo permite, además, abordar el análisis de filtros multicapa y filtros hechos con fibras de difracción de Bragg (que se estudiarán en un capítulo posterior). A modo de ejemplo, analizaremos el filtro de cavidad Fabry–Pérot. También estudiaremos los láseres de cavidad vertical ya que, aunque no se utilizan en comunicaciones ópticas, tienen costes de producción muy bajos. Estos dispositivos se usan hoy en día fundamentalmente en reproductores y grabadores de DVD y quizás, en el futuro, se empleen también en las comunicaciones por fibra óptica. Por último describiremos superficialmente el funcionamiento de los láseres de Bragg (DBR) cuya característica distintiva es que son láseres sintonizables en los que se puede controlar tanto la potencia emitida, como la longitud de onda y la fase de la onda electromagnética que generan. Estos son los láseres utilizados en los sistemas DWDM.

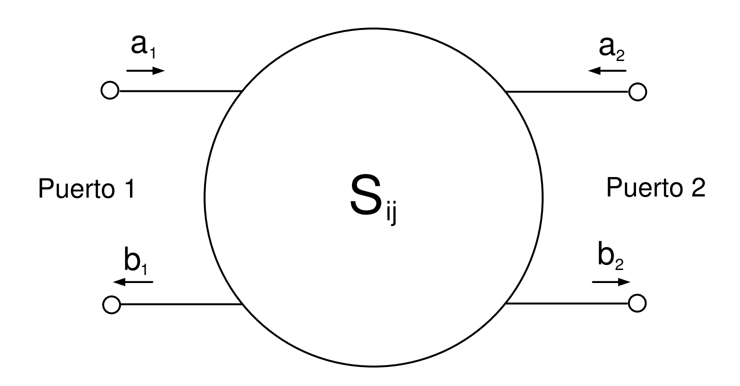


Figura 8.1: Esquema de una bipuerta descrita por una matriz de scattering.

8.1. Introducción

El análisis de muchos dispositivos, tanto ópticos como de microondas, puede abordarse mediante las llamadas matrices de scattering o matrices de reparto y otras formulaciones equivalentes [24]–[34]–[68]–[107]. Estas matrices pueden describir el funcionamiento de multipuertas (dispositivos con múltiples puertos de entrada y salida), aunque todos los dispositivos estudiados en este capítulo serán bipuertas, es decir, dispositivos de sólo dos puertas que se van a denotar por medio de los números "1" y "2". Las matrices de scattering describen las relaciones que impone el componente (o dispositivo) entre las ondas que *salen* de la bipuerta (ondas que denotaremos por medio de la letra "b") con respecto a las ondas que *entran* en ella (que se van a denotar con la letra "a"). Estas ondas no son ondas físicas, en el sentido de que no representan un campo eléctrico o magnético. Sus amplitudes son iguales a la raíz cuadrada del flujo de potencia (flujo entrante, si se trata de las ondas "a" y saliente si se trata de las ondas "b"), mientras que la fase es igual a la fase de cualquiera de los dos campos (eléctrico o magnético). En la fig. {8.1} se ha representado una bipuerta, cuyos parámetros S_{ij} relacionan las ondas "b" con las ondas "a" de la siguiente manera:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{8.1}$$

A continuación vamos a deducir la expresión de las matrices de scattering de los dos procesos físicos elementales (reflexión en una interfaz dieléctrica y propagación a lo largo de un medio dieléctricode longitud finita), a partir de los que podemos describir el funcionamiento de todos los dispositivos que vamos a estudiar a lo largo de este capítulo.

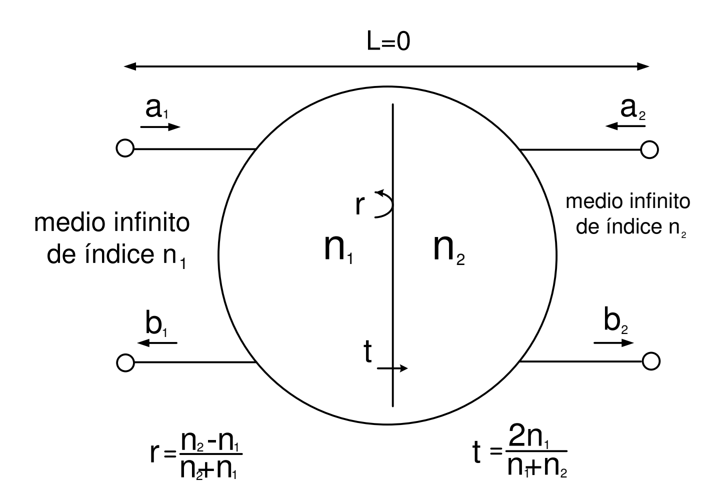


Figura 8.2: Esquema de una bipuerta correspondiente a una interfaz dieléctrica.

8.2. Matrices de scattering elementales

Empecemos con la interfaz dieléctrica, que no es más que una región de espesor infinitesimal (una superficie plana) en la que se produce una variación brusca del índice de refracción, que pasa de tomar el valor n_1 a la izquierda de esta superficie, a tomar otro valor distinto n_2 a la derecha, tal y como se representa en la fig. {8.2}. Supongamos que excitamos esta estructura desde el puerto "1" con una onda incidente a_1. Para garantizar que los puertos "1" y "2" no van a introducir reflexiones adicionales indeseadas, los "adaptamos", es decir, los continuamos con medios de longitud infinita cuyos índices toman los valores n_1 y n_2 respectivamente. De esta manera no se producen cambios de índice de refracción y no hay reflexiones en ninguno de los dos puertos. ¿Qué expresión analítica tiene la onda a_1? En primer lugar debe aparecer como término multiplicando la amplitud del campo eléctrico incidente $E_\circ$. A continuación debemos dividir por la raíz cuadrada de la impedancia del medio, $\sqrt{\frac{2\eta_\circ}{n_1}}$ para que el cuadrado de la onda represente el flujo de potencia incidente. Por último debe aparecer el término que indique que la propagación de la onda se produce de izquierda a derecha, $e^{-j\beta z}$, es decir:

$$a_1 = \frac{E_\circ \sqrt{n_1}}{2\eta_\circ} e^{-j\beta z} \tag{8.2}$$

¿Qué ondas se producen como respuesta ante esta onda incidente a_1? En primer lugar aparece una onda *transmitida* b_2 que obtenemos de la siguiente manera: su amplitud $E_\circ^{'}$ es el producto del coeficiente de transmisión en campo t multiplicado por la amplitud del campo incidente $E_\circ$, $E_\circ^{'} = tE_\circ$. A continuación debemos dividir por la raíz cuadrada de la impedancia del medio, $\sqrt{\frac{2\eta_\circ}{n_2}}$ para que el cuadrado de la onda represente el flujo de potencia transmitida. Por último debe aparecer el término que indica que la propagación de la onda se produce de izquierda a derecha, $e^{-j\beta z}$ [1], es decir:

[1]No estamos teniendo en cuenta que las constantes de propagación a ambos lados de la interfaz son ligeramente distintas.

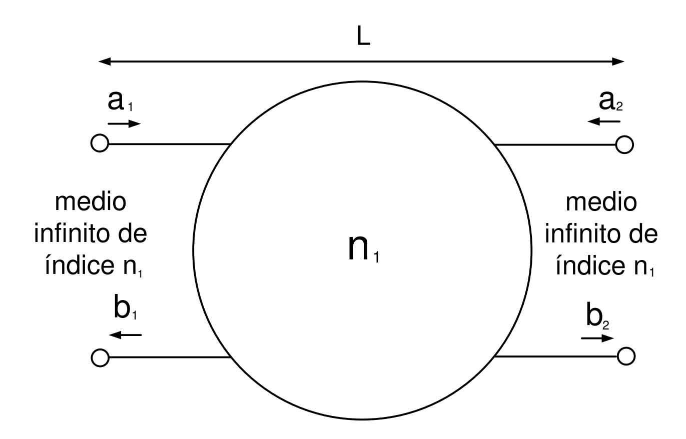


Figura 8.3: Esquema de una bipuerta correspondiente a un medio dieléctrico de longitud L.

$$b_2 = \frac{tE_\circ\sqrt{n_2}}{2\eta_\circ}e^{-j\beta z} \tag{8.3}$$

En segundo lugar aparece una onda *reflejada* b_1 que obtenemos de la siguiente forma: su amplitud $E_\circ^{''}$ es el producto del coeficiente de reflexión en campo r multiplicado por la amplitud del campo incidente $E_\circ$, $E_\circ^{''} = rE_\circ$. A continuación debemos dividir por la raíz cuadrada de la impedancia del medio, $\sqrt{\frac{2\eta_\circ}{n_1}}$ para que el cuadrado de la onda represente el flujo de potencia reflejada. Por último debe aparecer el término que indique que la propagación de la onda se produce de derecha a izquierda, $e^{j\beta z}$, es decir:

$$b_1 = \frac{rE_\circ\sqrt{n_1}}{2\eta_\circ}e^{j\beta z} \tag{8.4}$$

Ahora vamos a comprobar que se cumple la conservación del flujo de potencia, es decir, comprobemos que $a_1a_1^* = b_1b_1^* + b_2b_2^*$:

$$\begin{aligned} a_1a_1^* &= \frac{E_\circ^2 n_1}{2\eta_\circ} \\ b_1b_1^* + b_2b_2^* &= \frac{E_\circ^2 4n_1^2 n_2}{2\eta_\circ(n_1+n_2)^2} + \frac{E_\circ^2 n_1(n_2-n_1)^2}{2\eta_\circ(n_1+n_2)^2} = \frac{E_\circ^2 n_1}{2\eta_\circ} \end{aligned} \tag{8.5}$$

Si ensayamos el dispositivo por el puerto "2" llegamos a unos resultados análogos y deducimos que la bipuerta se puede describir por medio de la siguiente matriz de scattering:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} r & t_1 \\ t_2 & r \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{8.6}$$

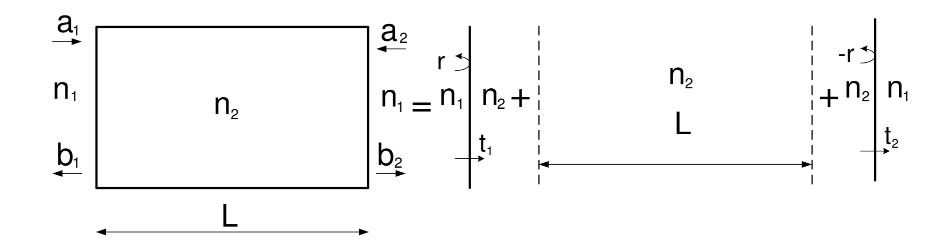


Figura 8.4: Cavidad Fabry-Pérot descompuesta en sus elementos fundamentales.

donde:

$$
\begin{aligned}
r &= \frac{n_1 - n_2}{n_1 + n_2} \\
t_1 &= \frac{2n_1}{n_1 + n_2} \\
t_2 &= \frac{2n_2}{n_1 + n_2}
\end{aligned} \tag{8.7}
$$

Vamos a estudiar a continuación la bipuerta (y su matriz de scattering asociada) correspondiente a la propagación por un medio dieléctrico de índice de refracción constante de valor n_1 y de longitud L. De nuevo, nos aseguraremos de que las puertas de entrada y salida ("1" y "2") no introducen reflexiones indeseadas "adaptando" impedancias, es decir, continuándolas con un medio de dimensiones infinitas cuyo índice de refracción sea constante e igual a n_1, tal y como se muestra en la fig. {8.3}.

Si excitamos el dispositivo por el puerto "1" con una onda incidente de expresión:

$$
a_1 = \frac{E_\circ \sqrt{n_1}}{\sqrt{2\eta_\circ}} e^{-j\beta z} \tag{8.8}
$$

resulta que esta onda se propaga sin reflexión hasta el puerto "2", por lo que se tiene:

$$
\begin{aligned}
b_1 &= 0 \\
b_2 &= \frac{E_\circ \sqrt{n_1}}{\sqrt{2\eta_\circ}} e^{-j\beta L} e^{-j\beta z}
\end{aligned} \tag{8.9}
$$

De igual forma, si excitamos el dispositivo por el puerto "2" obtenemos un resultado análogo: no se produce onda reflejada alguna y la onda sale por el terminal "1" con un desfase

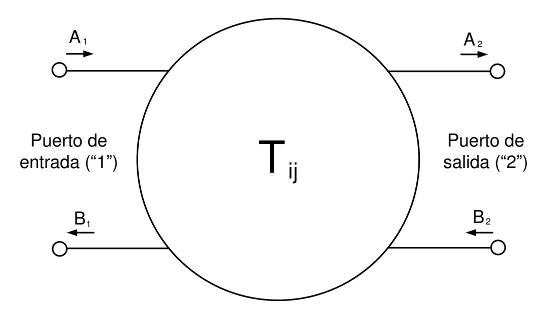


Figura 8.5: Descripción de bipuertas según el formalismo de las matrices de transmisión.

igual a $e^{-j\beta L}$. Deducimos que la matriz de scattering que describe el comportamiento de esta bipuerta es:

$$S_{ij} = \begin{bmatrix} 0 & e^{-j\beta L} \\ e^{-j\beta L} & 0 \end{bmatrix} \tag{8.10}$$

8.3. Cavidad resonante Fabry-Pérot

En la fig. {8.4} se ha representado una cavidad Fabry-Pérot. Ésta puede describirse como la sucesión de un interfaz dieléctrica caracterizada por unos coeficientes de reflexión y transmisión en campo r y t_1, la propagación a lo largo de un medio de índice de refracción constante n_2 y longitud L y, por último, otra interfaz dieléctrica cuyos coeficientes de reflexión y transmisión en campo son $-r$ y t_2. Ahora nos gustaría encontrar la matriz de scattering asociada a la cavidad a partir de las expresiones de las matrices de scattering de cada uno de sus elementos. Sin embargo el formalismo de matrices de scattering no está bien adaptado para la descripción de un conjunto de componentes conectados entre sí en cascada. El formalismo matricial que se adapta bien a este tipo de problemas es el de las matrices de transmisión, en el que se consideran como variables independientes las ondas en el puerto de salida y como variables dependientes las ondas en el puerto de entrada, además se clasifican las ondas según se propaguen por el dispositivo de izquierda a derecha (ondas "A") o de derecha a izquierda (ondas "B"). Esto se muestra en la fig. {8.5}.

Puesto que la matriz de tansferencia relaciona las ondas a la entrada de la bipuerta (puerto "1") con las ondas a la salida (puerto "2"), resulta que la matriz que representa a un conjunto de elementos colocados en serie, no es más que el producto ordenado de las matrices que representan a cada uno de estos elementos. Al igual que en el caso de las matrices de scattering, se supone que en los puertos de entrada y salida se adaptan impedancias a la hora de medir los coeficientes de la matriz de transmisión. También parece claro que, conociendo la descripción de una bipuerta mediante la matriz de scattering, podemos determinar la matriz de transmisión asociada y viceversa [2]. Las expresiones que relacionan estas dos descripciones

[2]Siempre que ambas sean no singulares.

son:

$$S_{ij} = \frac{1}{T_{11}}\begin{bmatrix} T_{21} & det(T) \\ 1 & -T_{12} \end{bmatrix}$$
$$T_{ij} = \frac{1}{S_{21}}\begin{bmatrix} 1 & -S_{22} \\ S_{11} & -det(S) \end{bmatrix} \quad (8.11)$$

Teniendo en cuenta estas relaciones, podemos escribir:

$$\left. T_{ij}^{(1)} = \begin{bmatrix} \frac{1}{t_1} & \frac{r}{t_1} \\ \frac{r}{t_1} & \frac{1}{t_1} \end{bmatrix} \; T_{ij}^{(2)} = \begin{bmatrix} e^{j\beta L} & 0 \\ 0 & e^{j\beta L} \end{bmatrix} \; T_{ij}^{(3)} = \begin{bmatrix} \frac{1}{t_2} & \frac{-r}{t_2} \\ \frac{-r}{t_2} & \frac{1}{t_2} \end{bmatrix} \right\} \Rightarrow$$

$$T_{ij}^{(1)}T_{ij}^{(2)}T_{ij}^{(3)} = \begin{bmatrix} \frac{e^{j\beta L} - r^2 e^{-j\beta L}}{t_1 t_2} & \frac{-re^{j\beta L} + re^{-j\beta L}}{t_1 t_2} \\ \frac{re^{j\beta L} - re^{-j\beta L}}{t_1 t_2} & \frac{e^{j\beta L} - r^2 e^{-j\beta L}}{t_1 t_2} \end{bmatrix} \quad (8.12)$$

Si ahora tenemos en cuenta (8.11), resulta que la matriz de scattering asociada a esta matriz de transmisión es:

$$S_{ij} = \begin{bmatrix} r - \frac{rt_1t_2e^{-2j\beta L}}{1 - r^2e^{-2j\beta L}} & \frac{t_1t_2e^{-j\beta L}}{1 - r^2e^{-2j\beta L}} \\ \frac{t_1t_2e^{-j\beta L}}{1 - r^2e^{-2j\beta L}} & r - \frac{rt_1t_2e^{-2j\beta L}}{1 - r^2e^{-2j\beta L}} \end{bmatrix} \quad (8.13)$$

Hasta ahora hemos dado por supuesto que la cavidad es pasiva, es decir, que $\beta \in \mathcal{R}$. Consideremos, por un instante, que la cavidad es activa, es decir, que $\beta' = \beta + j\frac{\alpha}{2}$ y tratemos de ver cuál es la condición de oscilación del láser de cavidad Fabry–Pérot resultante. Evidentemente, para que este dispositivo emita luz sin necesidad de que haya luz incidente, es necesario que los elementos de la matriz S_{ij} tengan un polo, para lo cual es necesario y suficiente que: $1 - r^2e^{-2j\beta' L} = 0$ de donde deducimos el siguiente par de relaciones (que son las condiciones de oscilación de un láser de cavidad Fabry-Pérot, que ya habíamos deducido en el capítulo correspondiente):

$$\frac{\alpha}{2} = \frac{1}{2L}Ln(\frac{1}{|r|^2})$$
$$\lambda_p = \frac{2Ln_2}{p} \quad (8.14)$$

Volvamos al caso de una cavidad sin ganancia (cavidad fría). En este caso lo que tenemos es un filtro de cavidad Fabry–Pérot. Para analizar las características de este filtro, vamos a representar los números complejos complejos S_{11} y S_{12}, dando sus módulos elevados al cuadrado y sus argumentos frente a la frecuencia de la radiación incidente, y parametrizándolos con respecto a distintos coeficientes de reflexión en campo r de la interfaz dieléctrica. Este resultado lo utilizaremos en el capítulo dedicado a los filtros ópticos y, más específicamente, cuando estudiemos los filtros sintonizables en transmisión. En la fig. {8.6} se representan los coeficiente de reflexión y transmisión complejos ($R = |S_{11}|^2$, $\angle S_{11}$) y ($T = |S_{12}|^2$, $\angle S_{12}$) frente a $\frac{f-f_\circ}{FSR}$, donde $f_\circ$ es la frecuencia central de trabajo (que ha de ser un múltiplo entero de FSR) y FSR es el denominado "rango espectral libre", cuyo valor viene dado por $FSR := \frac{c_\circ}{2n_2L}$.

Si definimos el coeficiente de reflexión en potencia de la interfaz dieléctrica como $R = |r|^2$, tendremos que la función de transferencia del filtro en transmisión viene definida por [104]:

$$T(f) := |S_{21}|^2 = \frac{(1-R)^2}{(1-R)^2 + 4R\sin^2(\frac{\pi f}{FSR})} \quad FSR := \frac{c_\circ}{2n_2L} \tag{8.15}$$

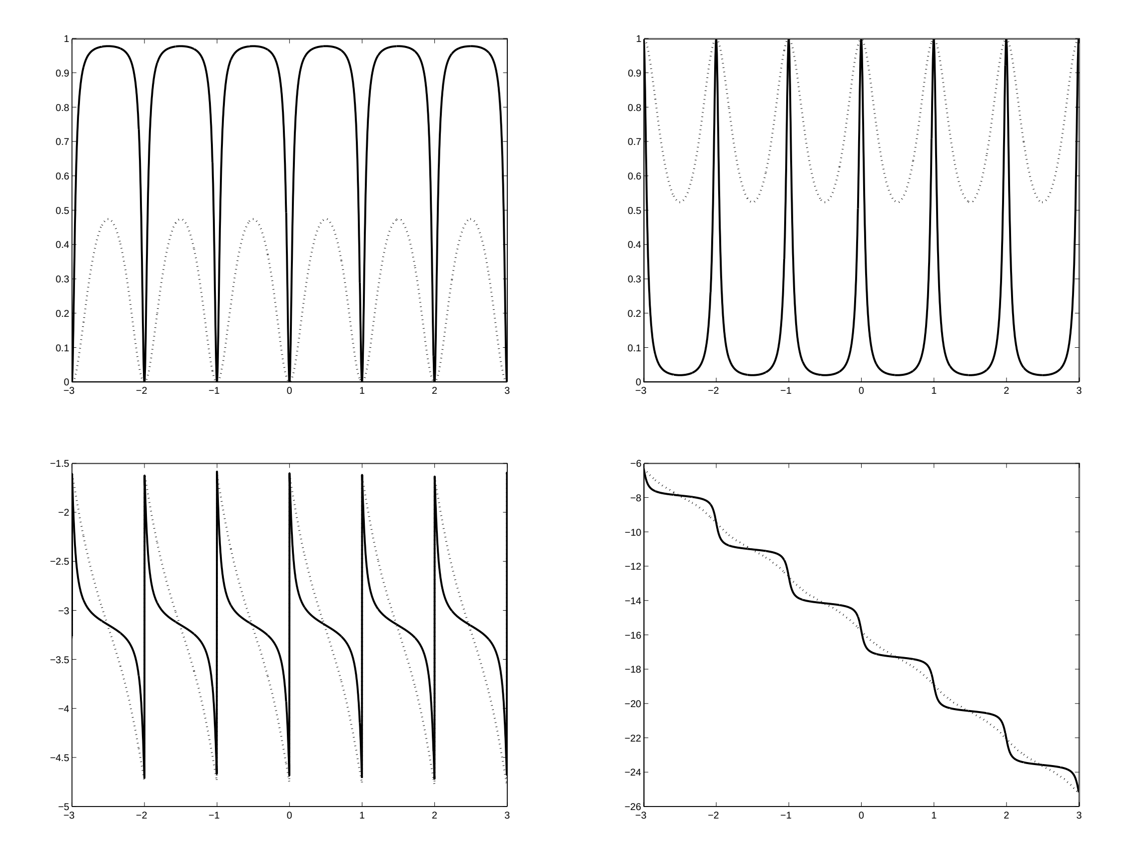


Figura 8.6: Coeficiente de reflexión (izquierda) y transmisión (derecha) complejos (módulo arriba y argumento abajo) de una cavidad Fabry–Pérot de $L = 300\mu m$, $r_1 = 0,4$ (trazo discontinuo) y $r_2 = 0,87$ (trazo continuo). Simulación realizada por Alejandro Carballar Rincón.

Es inmediato comprobar que el coeficiente de transmisión se hace máximo, $T(f) = T_{max} = 1$, para las frecuencias que son múltiplos enteros del rango espectral libre, $f = k \times FSR$ con $k = 1, 2, 3...$, mientras que es mínimo, $T(f) = T_{min} = (\frac{1-R}{1+R})^2$, para las frecuencias que son múltiplos semienteros del rango espectral libre, $f = \frac{k}{2} \times FSR$ con $k = 1, 3, 5....$ Vemos, pues, que la cavidad resonante Fabry–Pérot se comporta como un filtro paso de banda en transmisión periódico en frecuencia, siendo la anchura del periodo fundamental igual al rango espectral libre. También se puede comprobar que el ancho de banda del filtro, entendido como el intervalo de frecuencias para el que $T(f) \geq \frac{1}{2}$, es:

$$FWHM := \frac{c_\circ}{\pi n_2 L} \sin^{-1}(\frac{1-R}{2\sqrt{R}}) \tag{8.16}$$

Las cavidades resonantes pasivas Fabry–Pérot se utilizan mucho en comunicaciones ópticas como filtros paso–banda sintonizables en transmisión, fundamentalmente en sistemas densos de multiplexación en longitud de onda (DWDM).

8.4. Reflectores distribuidos

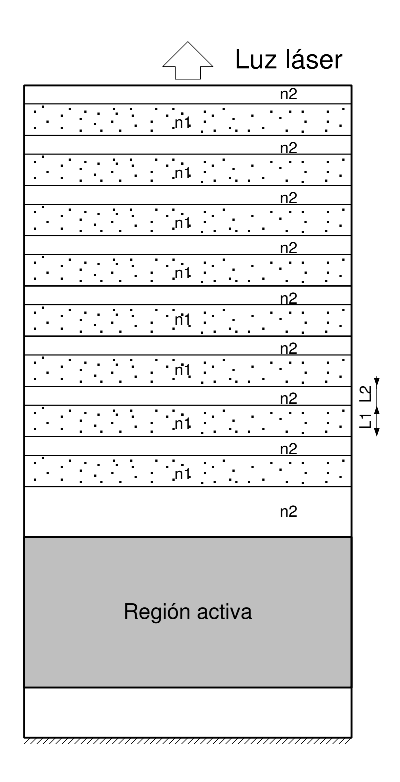


Figura 8.7: Láser de cavidad vertical (VCSEL).

Existen distintos tipos de láseres semiconductores de reflectores distribuidos. Vamos a considerar primero el denominado *láser de cavidad vertical* o VCSEL [8] que, aunque no se utiliza mucho en comunicaciones ópticas (están empezando a emplearse en algunos enlaces de primera ventana [58]), tiene un principio de funcionamiento fácil de entender.

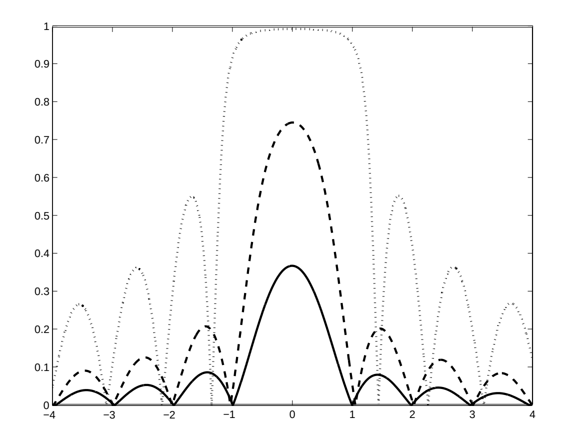
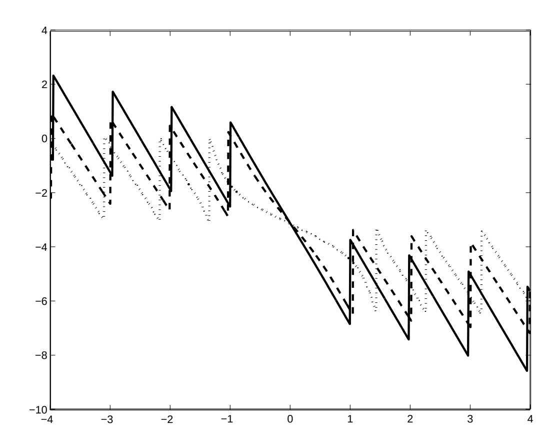



Figura 8.8: Coeficiente de reflexión complejo (módulo a la izquierda y argumento a la derecha) de la cavidad distribuida de un VCSEL de 8 (trazo continuo), 20 (trazo discontinuo) y 80 capas (trazo punteado). La variable independiente es $\frac{\delta L_g}{\pi}$. Simulación realizada por Alejandro Carballar Rincón.

Su estructura se muestra en la fig. {8.7}. En este tipo de láser se ha sustituido uno de los espejos de la cavidad Fabry–Pérot por un conjunto de capas de material dieléctrico cuyas anchuras e índices de refracción toman, alternativamente, dos valores distintos. Si el número de capas es suficientemente grande, el coeficiente de reflexión de la estructura es muy alto en un entorno estrecho de longitudes de onda centrado alrededor la denominada ”longitud de onda de Bragg”.

Queremos determinar la matriz de transmisión del reflector distribuido. Para ello nos fijamos en el periodo elemental que se repite a lo largo de la estructura: está formado por una capa de dieléctrico de espesor L_2 e índice de refracción n_2 seguido de otra capa, también de material dieléctrico, de espesor L_1 e índice de refracción n_1. Debemos hallar la matriz de transmisión de esta estructura, elevarla a la enésima potencia (el reflector está compuesto por un número n de estas estructura) y determinar, así, la matriz de transmisión del conjunto. Vemos que lo primero que encontramos es una interfaz dieléctrica asociada al cambio de índice del valor n_1 al valor n_2, a continuación nos encontramos con un fenómeno de propagación a lo largo de un medio de longitud L_2 e índice n_2, luego otra interfaz dieléctrica en la que el cambio de índice es desde el valor n_2 al valor n_1 y, por último, otra propagación a lo largo un medio de longitud L_1 e índice n_1. Para reducir el número de parámetros de los que depende el problema (para hacer la simulación numérica no es necesaria esta aproximación), vamos a suponer que los caminos ópticos asociados a las dos propagaciones son iguales: $L_{opt} = n_1 L_1 = n_2 L_2 \simeq \frac{n_1+n_2}{2}\frac{L_1+L_2}{2} \equiv \overline{n}\frac{\Lambda}{2}$. La penúltima igualdad es aproximadamente válida si los índices son muy parecidos entre sí, lo cual es siempre cierto (se consiguen reflectividades grandes apilando muchas capas sucesivas, no porque la reflectividad de una interfaz sea alta). Las matrices asociadas son:

$$T_{ij}^{(1)} = \begin{bmatrix} \dfrac{1}{t_1} & \dfrac{r}{t_1} \\ \dfrac{r}{t_1} & \dfrac{1}{t_1} \end{bmatrix} T_{ij}^{(2)} = \begin{bmatrix} e^{j\beta_\lambda \overline{n}\frac{\Lambda}{2}} & 0 \\ 0 & e^{-j\beta_\lambda \overline{n}\frac{\Lambda}{2}} \end{bmatrix} T_{ij}^{(3)} = \begin{bmatrix} \dfrac{1}{t_2} & \dfrac{-r}{t_2} \\ \dfrac{-r}{t_2} & \dfrac{1}{t_2} \end{bmatrix} T_{ij}^{(4)} = \begin{bmatrix} e^{j\beta_\lambda \overline{n}\frac{\Lambda}{2}} & 0 \\ 0 & e^{-j\beta_\lambda \overline{n}\frac{\Lambda}{2}} \end{bmatrix}$$

$$T_{ij}^{(1)}T_{ij}^{(2)}T_{ij}^{(3)}T_{ij}^{(4)} = \begin{bmatrix} \dfrac{e^{j\beta_\lambda \overline{n}\Lambda} - r^2}{t_1 t_2} & \dfrac{-re^{j\beta_\lambda \overline{n}\Lambda} + r}{t_1 t_2} \\ \dfrac{r - re^{-j\beta_\lambda \overline{n}\Lambda}}{t_1 t_2} & \dfrac{e^{-j\beta_\lambda \overline{n}\Lambda} - r^2}{t_1 t_2} \end{bmatrix} \tag{8.17}$$

En las expresiones anteriores tenemos que $\Lambda = L_1 + L_2$, $\overline{n} = \frac{n_1+n_2}{2}$, $r = abs(\frac{n_1-n_2}{n_1+n_2})$, $t_1 = \frac{2n_1}{n_1+n_2}$, $t_2 = \frac{2n_2}{n_1+n_2}$ y $\beta_\lambda = \frac{2\pi}{\lambda}$. Por supuesto, λ representa la longitud de onda de la luz en el vacío. Las operaciones matriciales hay que realizarlas numéricamente utilizando un programa de cálculo numérico, como, por ejemplo, MATLAB. En el caso de que el coeficiente de reflexión resultante sea inferior a 0.15 se pueden despreciar las reflexiones internas dentro de la estructura y se puede demostrar que el módulo del coeficiente de reflexión está descrito, aproximadamente, por la siguiente expresión [60]:

$$|S_{11}| \simeq 2mr\frac{\sin(\delta L_g)}{\delta L_g}$$
$$L_g := m\Lambda; \quad \delta := \frac{2\pi}{\lambda}\overline{n} - \frac{\pi}{\Lambda} \tag{8.18}$$

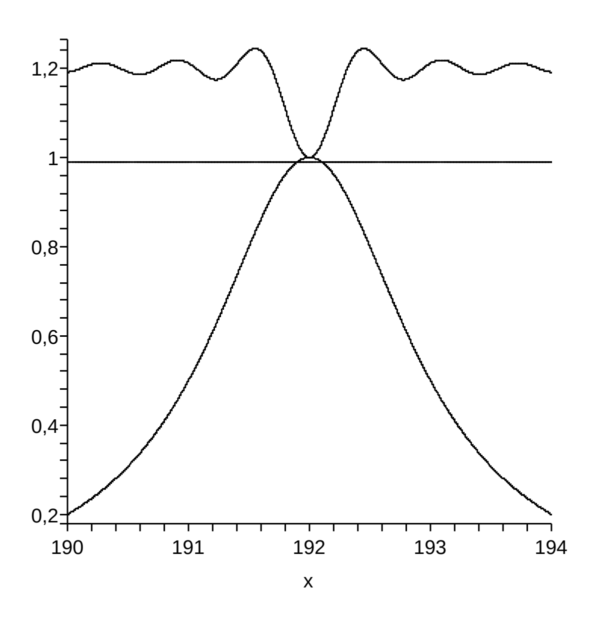


Figura 8.9: Curvas de ganancia y pérdida en un láser de cavidad distribuida. La línea horizontal representa la curva de pérdidas de un láser de cavidad Fabry–Pérot. La letra x representa la frecuencia en THz.

Donde m representa el número de periodos de la red y $2mr$ el coeficiente de reflexión máximo de la estructura (suponiendo que sea inferior a 0.15). En la fig. {8.8} se muestra el coeficiente de reflexión simulado en MATLAB para $2mr = 0{,}4$, $2mr = 1$ y $2mr = 4$ (para los cuales la aproximación anterior no es válida). Las condiciones de simulación son $\Lambda = 221{,}4nm$, $n_1 = 3{,}675$ y $n_2 = 3{,}5$. Observamos que a medida que mr va creciendo el

coeficiente de reflexión de la estructura toma un valor cada vez más cercano a la unidad para un intervalo de longitudes de onda cada vez mayor, mientras que en la característica de fase comprobamos que ésta va haciéndose cada vez más no lineal. La condición de reflexión máxima se produce, evidentemente, a la longitud de onda $\lambda_B = 2\overline{n}\Lambda$ que recibe el nombre de longitud de onda de Bragg. Dejamos como ejercicio para el lector el comprobar que a esta longitud de onda todas las reflexiones parciales se suman en fase.

8.5. Pérdidas en un láser de cavidad distribuida

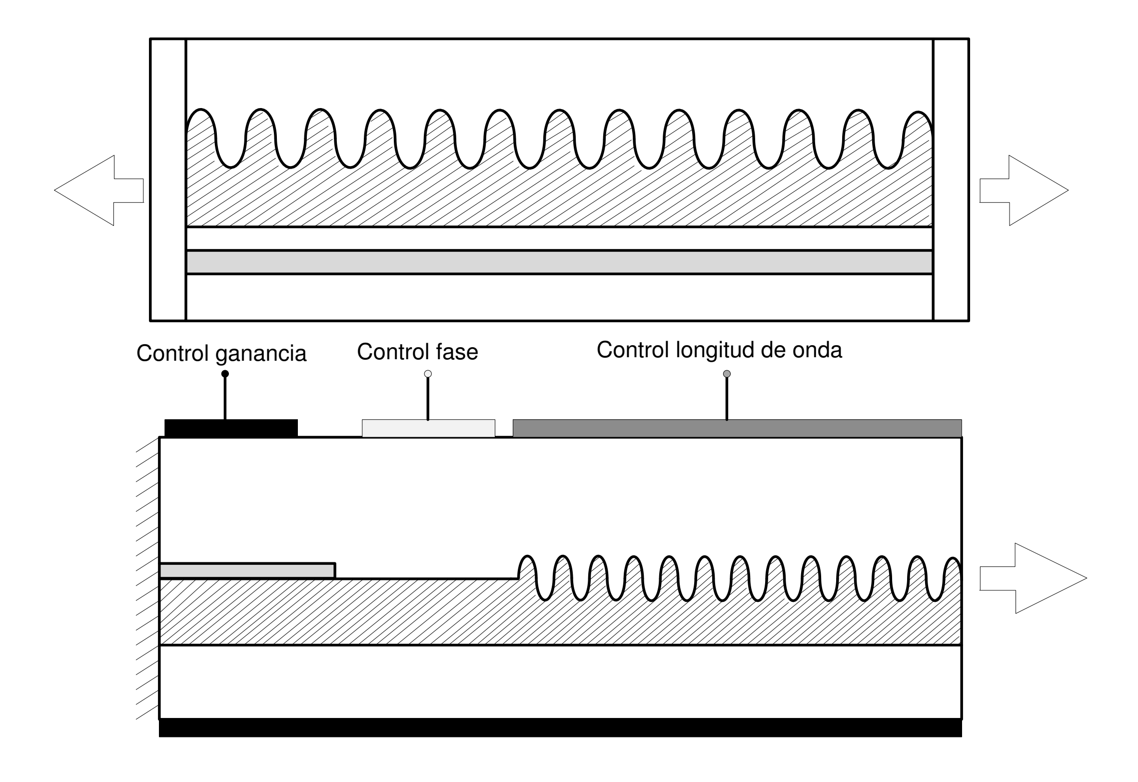


Figura 8.10: Secciones longitudinales de un DFB (arriba) y un DBR (abajo).

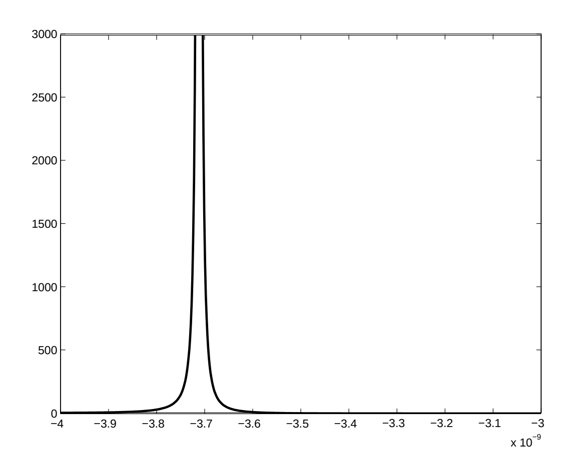
Los láseres de cavidad distribuida [8]–[56]–[58]–[74] que sistemáticamente se emplean en comunicaciones ópticas son el DFB (que emite a una longitud de onda fija) y el DBR (que emite a una longitud de onda programable dentro de un cierto rango). Ambos son láseres que emiten en un único modo longitudinal y por tanto son mucho más monocromáticos que el de cavidad Fabry–Pérot (que emiten potencia en varios modos longitudinales). Esto podemos entenderlo intuitivamente observando la fig. {8.9}. Vemos que la curva de pérdidas de una cavidad distribuida puede ser tangente a la curva de ganancia en un único punto con más facilidad que la curva de pérdidas de una cavidad Fabry–Pérot, ya que ésta es una recta horizontal.

Las cavidades resonantes de los láseres DFB y DBR son radicalmente distintas a la de los láseres VCSEL. En ellas se produce una variación longitudinal del índice *efectivo* que, recordemos, se define como $\frac{\beta}{k_o}$, donde β es la constante de propagación del modo guiado. Las variaciones senoidales del espesor de la capa semiconductora rayada de la fig. {8.10}, provocan una variación senoidal del índice efectivo a lo largo de la estructura del láser.

Estas perturbaciones senoidales son mucho más pequeñas que las del VCSEL (donde la luz atraviesa materiales de índices distintos) pero aumentando el número de periodos se consigue que el reflector distribuido tenga una reflectividad alta.

En el DFB normalmente [3] sólo se puede controlar un parámetro, la potencia óptica, regulando la corriente eléctrica que inyectamos en la región activa. En el DBR se pueden, sin embargo, controlar tres parámetros de funcionamiento distintos [8]–[56]: la potencia óptica, la frecuencia de emisión (dentro de un determinado rango, claro está) y, por último, la fase del campo electromagnético emitido. La potencia se controla regulando la corriente que se inyecta en la región activa, la longitud de onda se puede modificar en un rango amplio inyectando una corriente en la zona correspondiente al reflector distribuido. De esta manera se produce una variación en la concentración de portadores que modifica el índice de refracción medio de esta zona. Este cambio de índice produce una variación en la longitud de onda de Bragg del reflector distribuido. Por último, la fase puede controlarse también inyectando una corriente ajustable en la tercera región del DBR (cambiando, así, el índice del medio). Una sección longitudinal de un DFB bimodal y un DBR se muestra en la fig. {8.10}.

[3] Existen DFBs cuyas estructuras son distintas a las vistas hasta ahora, en los que se puede modificar ligeramente por efecto térmico la longitud de onda de emisión.



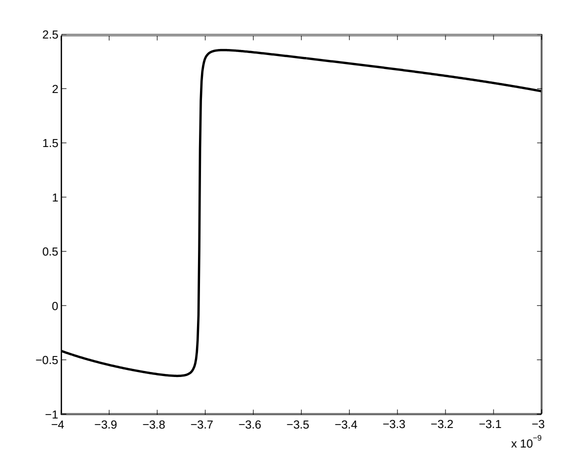


Figura 8.11: Coeficiente de transmisión complejo del DBR (módulo elevado al cuadrado a la izquierda y fase a la derecha). La longitud de onda de emisión es $1296,3nm$. La variable independiente es $\lambda - lambda_B$. Simulación realizada por Alejandro Carballar Rincón.

8.6. Problemas

Problema 10: Simular el comportamiento de un láser DBR utilizando el formalismo de las matrices de transmisión. Sustituya la perturbación senoidal del reflector de Bragg por una cuadrada cuyo primer armónico coincida con la amplitud de la perturbación senoidal. Suponga que sólo hay ganancia en la zona de la cavidad donde el índice medio es constante. La longitud de onda de Bragg es $\lambda_B = 1300nm$, el índice de refracción medio vale $n_\circ = 3{,}5$, la amplitud de la perturbación senoidal $n_1 = 0{,}0021$, $\Lambda = 185,714nm$ y la longitud del reflector es $L_2 = 125\mu m$. La longitud de la zona del índice sin perturbación senoidal es $L_1 = 125\mu m$, la ganancia de la región activa $n_2 = 0{,}00345$ y el coeficiente de reflexión del espejo: $R = 31\,\%$.

Solución: La forma de resolver este problema es calcular el coeficiente de transmisión de esta estructura, que está constituida por una interfase de coeficiente de reflexión R=31 % (aunque se podría haber dado un tratamiento a esta superficie, de forma que el coeficiente de reflexión en potencia fuese próximo a la unidad), una región dieléctrica de longitud $L = 125\mu m$ e índice $n = 3,5 - j0,00345$ (la parte imaginaria representa la ganancia del medio activo, que se ha hecho igual a la ganancia umbral del láser) y, por último, de un reflector de Bragg de longitud $125\mu m$ de índice $n = 3,5 + 0,0021\sin(\frac{2\pi}{\Lambda}z)$. Por tanto podemos calcular la matriz de transmisión del conjunto y determinar el coeficiente de transmisión o reflexión asociado. La ganancia del medio se ha tomado igual al valor umbral, de manera que el coeficiente de reflexión (o el de transmisión) se hace infinito a una determinada longitud de onda. Esa es, precisamente, la longitud de onda de emisión del láser.

En la {8.11} hemos representado el coeficiente de transmisión complejo del DBR (a la izquierda hemos representado el modulo elevado al cuadrado y a la derecha la fase). La longitud de onda nominal de emisión del DBR es $3,7nm$ inferior a la longitud de onda de Bragg, aunque no hemos considerando inyección de corriente alguna en la zona del reflector (no hemos simulado este efecto que podría cambiar la longitud de onda de emisión del láser).

Capítulo 9

Anchura espectral de un láser monomodo

Resumen: En este capítulo vamos a determinar la anchura espectral de emisión de un láser monomodo en condiciones estacionarias, es decir, emitiendo una potencia constante. Veremos que la emisión espontánea es la responsable de que, incluso en estas condiciones, se produzcan fluctuaciones aleatorias en la fase del campo electromagnético emitido por el dispositivo, lo que provoca un ensanchamiento del espectro emitido. Para reducirlo se puede aumentar la potencia de emisión, ya que el producto potencia emitida multiplicada por la anchura espectral es igual a una constante. Los láseres a semiconductor son intrínsecamente más ruidosos que los de niveles discretos, ya que en éstos se produce un acoplamiento entre los ruidos de amplitud y fase asociados a la emisión espontánea. Por este motivo estudiaremos primero el caso más simple de un láser de niveles discretos para, después, incluir los fenómenos de acoplamiento de las dos componentes de ruido y deducir las características espectrales de los láseres semiconductores.

9.1. Introducción

Nuestro objetivo es analizar el espectro de emisión de un láser monomodo semiconductor bajo condiciones de régimen permanente (es decir, emitiendo una potencia óptica constante). Puesto que este estudio es bastante complicado, preferimos determinar primero el espectro de emisión de un láser de niveles discretos y generalizar, después, el resultado para el caso de un láser semiconductor. En cualquier caso, la razón última por la que un láser monomodo cualquiera no emite un campo electromagnético completamente senoidal (en cuyo caso su densidad espectral de potencia sería una delta de Dirac) es que en la cavidad resonante hay fotones generados por emisión espontánea que producen fluctuaciones aleatorias en la amplitud y la fase del campo electromagnético emitido por el láser. Estas fluctuaciones producen, inevitablemente, un ensanchamiento espectral de la luz generada por el dispositivo.

Un análisis riguroso de cualquier fenómeno relacionado con la emisión espontánea obliga a cuantificar la radiación electromagnética. Para evitar esto y limitarnos a un análisis semiclásico del problema, vamos a introducir un postulado (que, evidentemente, es indemostrable) concerniente a la emisión espontánea: *el efecto de la emisión espontánea en un conjunto de átomos que amplifican radiación electromagnética se reduce a añadir a los fotones de esta radiación n_{sp} fotones incorrelados por modo* [92]. El término n_{sp} se conoce con el nombre coeficiente de emisión espontánea en exceso y, para los láseres de niveles discretos, tiene la expresión $n_{sp} = \frac{N_2}{N_2 - N_1}$. Si el medio amplificador es un semiconductor, entonces el coeficiente de emisión espontánea en exceso viene dado por $n_{sp} \simeq \frac{1}{1-e^{-\frac{F_c - F_v - E_g}{kT}}}$.

En los siguientes apartados (salvo el último) vamos a ir describiendo, sucesivamente, los distintos términos que introducen o sacan fotones de la cavidad resonante de un láser de *niveles discretos*. En el segundo apartado, estudiaremos una cavidad fría en la que únicamente salen fotones por pérdidas asociadas a las reflexiones parciales en los espejos y al scattering Rayleigh. Comprobaremos que la densidad espectral de energía de la luz emitida es inversamente proporcional al tiempo medio de permanencia del fotón en la cavidad. En el tercero, tendremos en cuenta, además, la ganancia óptica del medio pero no los fotones generados por emisión espontánea. En estas condiciones (imposibles de conseguir en la práctica) el tiempo medio de vida del fotón se hace infinito, ya que la ganancia del medio (emisión estimulada menos absorción) repone los fotones que se pierden. En estas condiciones (que son irrealizables), la densidad espectral de potencia de la luz generada por el láser vendría dada por una delta de Dirac. En cuarto apartado incluiremos, también, el término de emisión espontánea y deduciremos la expresión real de la anchura espectral de emisión de un láser monomodo de niveles discretos. Por último, estudiaremos el acoplamiento entre los ruidos de amplitud y fase que se produce en los láseres semiconductores y veremos que el espectro de emisión es considerablemente más ancho que el de un láser de niveles discretos.

9.2. Cavidad fría

Consideremos la cavidad resonante de un láser en el que súbitamente interrumpimos el bombeo, de manera que solamente tenemos pérdidas, tal y como se muestra en la fig.

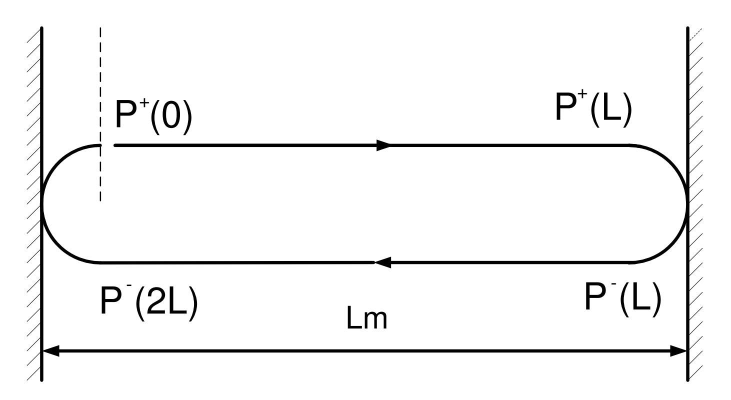


Figura 9.1: Cavidad fría.

{9.1}. La evolución del número de fotones en la cavidad, viene dada por:

$$\begin{aligned} &P^+(0) = P_\circ \\ &P^+(L) = P_\circ e^{-\alpha_{sc}L} \\ &P^-(2L) = RP_\circ e^{-\alpha_{sc}L} \\ &P^-(2L) = RP_\circ e^{-\alpha_{sc}2L} \\ &P^+(2L) = R^2 P_\circ e^{-\alpha_{sc}2L} \end{aligned} \tag{9.1}$$

Ahora vamos a distribuir las pérdidas localizadas en los espejos a lo largo de toda la cavidad, para lo que definimos un coeficiente de pérdidas por unidad de longitud asociado a las reflexiones parciales en éstos. Definiendo $\alpha_{esp} := \frac{1}{2L} Ln(\frac{1}{R^2})$ nos queda:

$$\begin{aligned} P^+(2L) = P_\circ e^{-(\alpha_{sc}+\alpha_{esp})2L} \Rightarrow P(x) = P_\circ e^{-(\alpha_{sc}+\alpha_{esp})x} \Rightarrow \\ \left(\frac{dP}{dx}\right)_{per} = \frac{dP}{dx}\frac{dx}{dt} = -2(\alpha_{sc}+\alpha_{esp})P\frac{c_\circ}{n} = -\frac{P}{\tau_p} \end{aligned} \tag{9.2}$$

donde la constante de tiempo de pérdidas de la cavidad fría (pérdidas que no están relacionadas con la atenuación del medio activo) viene dada por: $\tau_p := \frac{n}{2c_\circ(\alpha_{sc}+\alpha_{esp})}$ y donde $\alpha_{esp} := \frac{1}{2L} Ln(\frac{1}{R^2})$. El significado físico de τ_p es evidente: es el tiempo que, en promedio, tarda en salir un fotón de la cavidad por alguno de los dos mecanismos mencionados, que son o bien una reflexión parcial en uno de los espejos o bien un proceso de scattering Rayleigh.

Si ahora integramos la ecuación diferencial $\frac{dP}{dt} = -\frac{P}{\tau_p}$ queda: $P(t) = P_\circ e^{-\frac{P}{\tau_p}}$. A partir de esta expresión (evolución temporal de la potencia óptica presente en la cavidad del láser) podemos calcular fácilmente la evolución temporal del campo eléctrico:

$$E(t) = E_\circ e^{-\frac{t}{2\tau_p}} e^{j\omega_\circ t} \Rightarrow E(\omega) = \frac{E_\circ}{\frac{1}{2\tau_p} + j(\omega - \omega_\circ)} \tag{9.3}$$

La densidad espectral de energía no normalizada viene dada por:

$$S(\omega) = E^*(\omega)E(\omega) = \frac{E_\circ^2}{\frac{1}{4\tau_p^2} + (\omega - \omega_\circ)^2} \tag{9.4}$$

Para expresar esta densidad espectral de energía en forma normalizada, imponemos que el área definida por la función $S^{nor}(\omega)$ y el eje ω sea igual a 1.

$$\int_{-\infty}^{\infty} S(\omega)d\omega = \int_{-\infty}^{\infty} \frac{E_\circ^2}{\frac{1}{4\tau_p^2} + (\omega - \omega_\circ)^2} d\omega = \int_{-\infty}^{\infty} \frac{E_\circ^2 4\tau_p^2}{1 + [2\tau_p(\omega - \omega_\circ)]^2} d\omega \tag{9.5}$$

Si ahora definimos $x = 2\tau_p(\omega - \omega_\circ)$ y sustituimos en (9.5) nos queda:

$$\int_{-\infty}^{\infty} S(\omega)d\omega = E_\circ^2 2\tau_p \int_{-\infty}^{\infty} \frac{dx}{1 + x^2} = \pi E_\circ^2 2\tau_p \Rightarrow S^{nor}(\omega) = \frac{S(\omega)}{\pi E_\circ^2 2\tau_p} = \frac{\frac{2}{\pi\Delta\omega}}{1 + (\frac{\omega - \omega_\circ}{\Delta\omega/2})^2} \tag{9.6}$$

Donde $\Delta\omega = \frac{1}{\tau_p}$ es la anchura espectral de la energía del campo emitida por la cavidad fría. Vemos que cuanto mayor es el tiempo medio de permanencia de un fotón en la cavidad τ_p, menor es la anchura espectral $\Delta\omega$ de emisión y viceversa.

9.3. Cavidad caliente sin emisión espontánea

Ahora introducimos el término de ganancia en la ecuación de balance del número de fotones que hay en la cavidad (para simplificar la notación introducimos el factor de confinamiento en la definición de τ_g):

$$\frac{dP}{dt} = \frac{P}{\tau_g} - \frac{P}{\tau_p} = \frac{P}{\frac{\tau_g\tau_p}{\tau_p - \tau_g}} \tag{9.7}$$

En condiciones estacionarias todas las derivadas temporales son nulas. Como estamos suponiendo que el láser emite luz, resulta que $P \neq 0$ por lo que se cumple que $\tau_p = \tau_g$ y, por tanto, el tiempo medio de permanencia de un fotón en la cavidad resulta ser infinito con lo que la densidad espectral de potencia de la radiación emitida es:

$$S(\omega) = E_\circ^2\delta(\omega - \omega_\circ) \Rightarrow S^{nor}(\omega) = \delta(\omega - \omega_\circ) \tag{9.8}$$

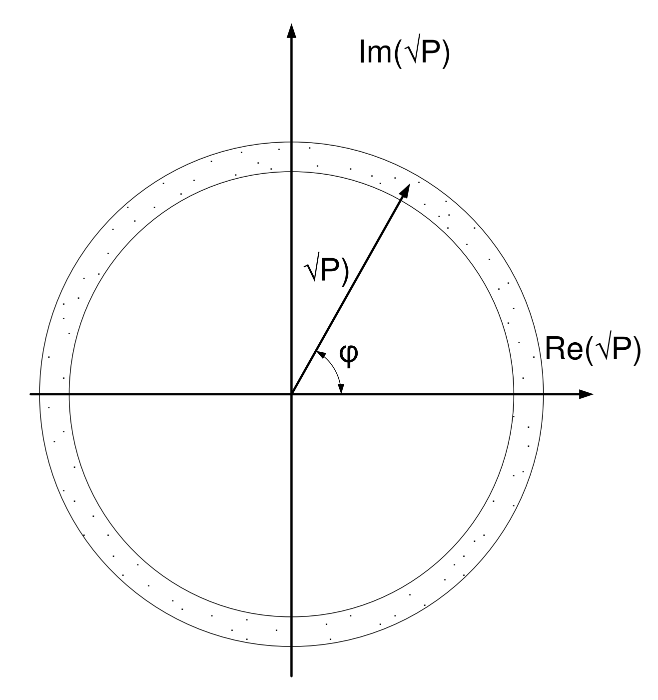


Figura 9.2: Campo electromagnético dentro de la cavidad.

9.4. Cavidad caliente con emisión espontánea

Si ahora incluimos el término de emisión espontánea, la ecuación de balance del número de fotones que hay dentro de la cavidad resonante es:

$$\frac{dP}{dt} = \frac{P}{\tau_g} - \frac{P}{\tau_p} + R \tag{9.9}$$

donde R es el número de fotones espontáneos que, por unidad de tiempo, se añaden a los generados por los procesos estimulados (emisión estimuladas menos absorción). Estos fotones pertenecen al mismo modo espacial y temporal que los estimulados, pero no tienen por qué estar en fase con la radiación que se está amplificando (de hecho se suman con un desfase arbitrario). Si imponemos condiciones de régimen permanente resulta:

$$0 = \frac{P}{\tau_g} - \frac{P}{\tau_p} + R \Rightarrow \frac{1}{\tau_p} - \frac{1}{\tau_g} = \frac{R}{P} \Rightarrow \tau_{eq} = \frac{\tau_p \tau_p}{\tau_p + \tau_g} = \frac{P}{R} \tag{9.10}$$

Vemos que ahora el tiempo medio de permanencia de un fotón dentro de la cavidad es finito e igual al cociente entre el número de fotones que hay en la cavidad resonante, P, y la tasa de emisión espontánea, R. Si la tasa de emisión espontánea fuese nula, de nuevo tendríamos un tiempo infinito de permanencia dentro de la cavidad. Vemos, pues, que el motivo por el que el láser tiene una anchura espectral de emisión no nula es, precisamente, la existencia de fotones que, generados por emisión espontánea, se suman a los producido por fenómenos estimulados. En la fig. {9.2} se ha representado el campo electromagnético dentro de la cavidad del láser (hemos eliminado, claro está, el término de pulsación $e^{j\omega_o t}$).

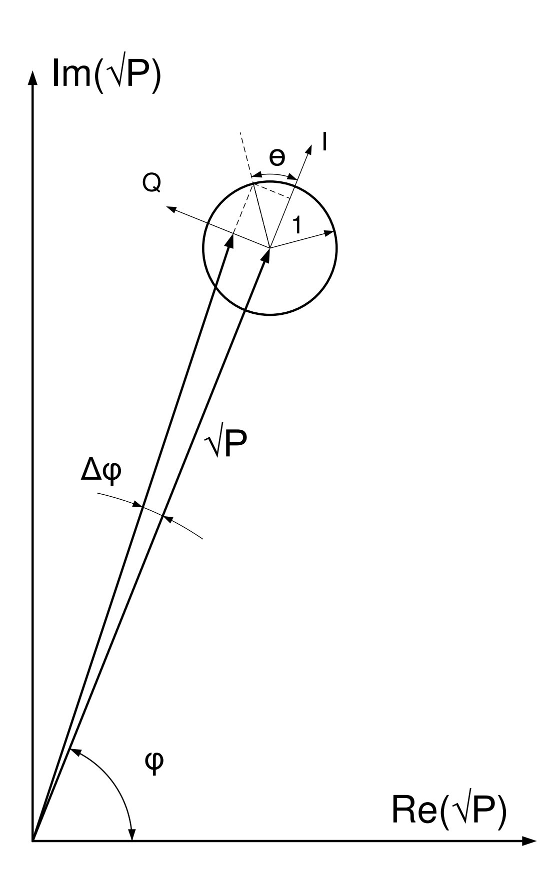


Figura 9.3: Variación de fase del campo electromagnético como consecuencia de que se suma un fotón generado por emisión espontánea.

Se observa que tanto la amplitud $\sqrt{P}$ como la fase φ van variando con el tiempo como consecuencia de los fotones añadidos por emisión espontánea. Estos fotones tiene una fase totalmente incorrelada con la fase del campo electromagnético que hay en la cavidad φ, de ahí que se produzcan variaciones tanto de amplitud como de fase. Las fluctuaciones de fase son las que provocan un mayor ensanchamiento espectral del láser y van a ser las únicas que vamos a considerar en lo que sigue.

En la fig. {9.3} hemos representado el cambio sufrido por el fasor campo eléctrico cuando se añade un fotón procedente de la emisión espontánea. La componente en fase (I) produce un cambio de la amplitud del campo eléctrico cuya influencia en el espectro de emisión del láser podemos considerar despreciable. Sin embargo, la componente en cuadratura (Q) produce una variación de fase que es la principal responsable del espectro de emisión del láser. Como la fase del fotón agregado por emisión espontánea puede ser cualquiera, esto significa que θ es una variable aleatoria con distribución uniforme en el intervalo $(0, 2\pi)$, es decir, $\theta \sim U(0, 2\pi)$, con lo que podemos escribir:

$$\mathrm{tg}(\Delta\varphi) = \frac{\sin(\theta)}{\sqrt{P}} \tag{9.11}$$

como $\varphi \ll 1$ tenemos que $\Delta\varphi \simeq \frac{\sin(\theta)}{\sqrt{P}}$ y resulta:

$$
\begin{aligned}
&< \Delta\varphi >= \frac{1}{2\pi}\int_0^{2\pi} \frac{\sin(\theta)}{\sqrt{P}} d\theta = 0 \\
&< \Delta\varphi^2 >= \frac{1}{2\pi}\int_0^{2\pi} \frac{\sin^2(\theta)}{\sqrt{P}} d\theta = \frac{1}{2\pi}\int_0^{2\pi} \frac{1-\cos(2\theta)}{2P} d\theta = \frac{1}{2P}
\end{aligned} \tag{9.12}
$$

En realidad tenemos una tasa R de fotones generados por unidad de tiempo, de forma que la variación de fase $\Delta\varphi$ por unidad de tiempo tiene como estadísticos principales:

$$
\begin{aligned}
&< \Delta\varphi >= 0 \\
&< \Delta\varphi^2 >= \frac{R}{2P}
\end{aligned} \tag{9.13}
$$

Pero, ¿qué valor toma R? Vimos en la introducción de este capítulo que para tener en cuenta los fenómenos de emisión espontánea había que tratarla como si fuese emisión estimulada, sustituyendo los P fotones que se tienen en la cavidad (generados por procesos estimulados, esto es, por emisión estimulada menos absorción) por n_{sp} fotones producidos por emisión espontánea. Dado que el número de fotones producidos por unidad de tiempo por fenómenos estimulados es $\frac{P}{\tau_g} \simeq \frac{P}{\tau_p}$ deducimos que $R \simeq \frac{n_{sp}}{\tau_p}$ (número de fotones generados por unidad de tiempo por procesos espontáneos) quedando:

$$
\left.\begin{aligned}
&< \Delta\varphi >= 0 \\
&< \Delta\varphi^2 >= \frac{n_{sp}}{2P\tau_p}
\end{aligned}\right\} \Rightarrow \left.\begin{aligned}
&< \Delta\varphi(\tau) >= 0 \\
&< \Delta\varphi^2(\tau) >= \frac{n_{sp}\tau}{2P\tau_p}
\end{aligned}\right\} \tag{9.14}
$$

Dado que el campo eléctrico que hay dentro de la cavidad es una señal aleatoria, para calcular el espectro de ésta tenemos que calcular su función de autocorrelación y, una vez obtenida, calcular su transformada de Fourier (teorema de Wiener–Kintchine [85]–[104]).

$$
E(t) = E_\circ e^{j[\omega_\circ t + \phi(t)]} \Rightarrow < E^*(t)E(t+\tau) >= E_\circ^2 e^{j\omega_\circ \tau} < e^{j[\phi(t+\tau)-\phi(\tau)]} > \tag{9.15}
$$

Ahora, sabiendo que $\Delta\varphi(\tau)$ se distribuye como una variable aleatoria gaussiana de media 0 y desviación típica $\sqrt{< \Delta\varphi^2 >}$, es decir, $\Delta\varphi \sim N(0; \sqrt{\frac{n_{sp}\tau}{\tau_p 2P}})$, se puede demostrar de forma rigurosa que la función de autocorrelación viene dada por (9.16). El desarrollo que se muestra a continuación no es más que una regla mnemotécnica para recordar el resultado:

$$
\begin{aligned}
&R_E(\tau) :=< E^*(t)E(t+\tau) >= E_\circ^2 e^{j\omega_\circ \tau} < e^{j[\phi(t+\tau)-\phi(\tau)]} >= \\
&= E_\circ^2 e^{j\omega_\circ \tau} < [1 + j(\Delta\varphi(\tau) - \tfrac{1}{2}\Delta\varphi^2(\tau) + ...] >= \\
&= E_\circ^2 e^{j\omega_\circ \tau}[1 + j < \Delta(\tau) > -\tfrac{1}{2} < \Delta\varphi^2(\tau) > + ...] = \\
&= E_\circ^2 e^{j\omega_\circ \tau}[1 - \tfrac{1}{2} < \Delta\varphi^2(\tau) > + ...] = E_\circ^2 e^{j\omega_\circ \tau} e^{-\frac{1}{2}<\Delta\varphi^2(\tau)>}
\end{aligned} \tag{9.16}
$$

Si ahora definimos $\Delta\omega_l := \frac{n_{sp}}{\tau_p 2P}$ nos queda:

P número de fotones en la cavidad

P/2 la mitad va hacia la derecha

c_o/n_g velocidad de los fotones

$\hbar\omega$ energía de un fotón

cavidad resonante

$P/2\ c_o/n_g\ \hbar\omega$ potencia x longitud

α_E pérdidas por unidad de l. del espejo

$W=P/2\ c_o/n_g \hbar\omega\alpha_E$

Potencia óptica que sale del espejo

Figura 9.4: Relación entre la potencia óptica emitida por el láser y el número de fotones que hay en la cavidad.

$$\begin{aligned} R_E(\tau) &= E_\circ^2 e^{-\frac{1}{2}\Delta\omega_l \tau} e^{j\omega_\circ \tau} \\ \Delta\omega_l &:= \frac{n_{sp}}{\tau_p 2P} \end{aligned} \tag{9.17}$$

Por tanto, teniendo en cuenta que una desviación típica es siempre una cantidad positiva, el espectro de emisión normalizado del láser viene dado por:

$$\left.\begin{aligned} E^{nor}(\omega) = \frac{S_E(\omega)}{E_\circ^2} := \frac{1}{E_\circ^2}\int_{-\infty}^{\infty} R_E(\tau) e^{-j\omega\tau} d\tau = \frac{\dfrac{2}{\pi\Delta\omega_l}}{1+\left(\dfrac{\omega-\omega_\circ}{\Delta\omega_l/2}\right)^2} \\ \Delta\omega_l := \frac{n_{sp}}{\tau_p 2P} \end{aligned}\right\} \tag{9.18}$$

Sin embargo esta expresión de la anchura espectral, $\Delta\omega_l = \frac{n_{sp}}{\tau_p 2P}$, no es fácil de interpretar ya que P representa el número de fotones que hay dentro de la cavidad. Nos gustaría relacionar P con W, la potencia óptica emitida por el láser, que es la variable física asociada a P que se puede medir con facilidad. Si observamos la fig. {9.4}, resulta claro que podemos escribir:

$$\left.\begin{aligned} W &= \frac{P}{2}\frac{c_\circ}{n_g}\hbar\omega\alpha_E \\ \tau_p &= \frac{n_g}{c_\circ(\alpha_{sc}+\alpha_E)} \end{aligned}\right\} \Rightarrow \Delta\omega_l = \frac{n_{sp}}{\tau_p 2P} = n_{sp}\frac{\hbar\omega\alpha_E(\alpha_E+\alpha_{sc})c_\circ^2}{4Wn_g^2} \tag{9.19}$$

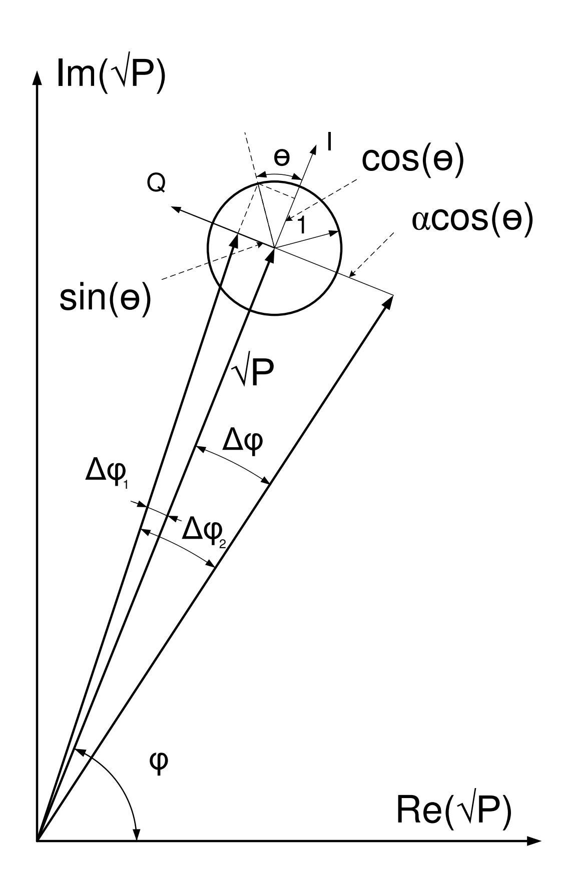


Figura 9.5: Variación de fase del campo electromagnético como consecuencia de que se suma un fotón generado por emisión espontánea y el acoplamiento entre el ruido de amplitud y el ruido de fase.

que es la famosa relación de Schawlow y Townes [91].

Si ahora tomamos valores típicos para n_{sp}, α_E y α_{sc} nos queda un valor $\Delta\omega_l W \simeq 2\pi \times 5MHz \times mW$ resultado que está de acuerdo con las medidas experimentales del espectro de emisión de un láser de niveles discretos. Sin embargo en un DFB o un DBR los valores experimentales típicos son $\Delta\omega_l W \simeq 2\pi \times 100MHz \times mW$. Esto quiere decir que hay algún fenómeno físico importante en los láseres a semiconductor que no hemos tenido en cuenta [51].

9.5. Laser semiconductor

El índice de refracción de un medio material es una magnitud compleja, $n = n_R - jn_I$, donde la parte real n_R afecta al desfase que sufre la onda electromagnética que se propaga por el medio (la constante de propagación en un medio dieléctrico es $\beta(\omega) = k_\circ n_R$), mientras que la parte imaginaria n_I es la responsable de la atenuación (o de la amplificación) que sufre el módulo (recordemos que el coeficiente de atenuación de un medio dieléctrico es $\frac{\alpha}{2} = k_\circ n_I$). A la parte imaginaria le hemos asignado un signo negativo porque estamos estudiando un láser y, por tanto, el medio material tiene ganancia, no atenuación. La parte real e imaginaria

del índice de refracción no pueden variar independientemente, ya que derivan de una misma función analítica compleja, $\chi(\omega)$. Recordemos que las relaciones integrales que ligan n_R y n_I se conocen con el nombre de relaciones de Kramers–Kronig. Una aproximación grosera de esta relación (válida para un ancho de banda espectral pequeño) es $\frac{\Delta n_R}{\Delta n_I} = a$, donde Δn_R y Δn_I son las variaciones de la parte real e imaginaria del índice de refracción producidas por la emisión espontánea y a es una constante positiva que recibe el nombre de *factor de ensanchamiento espectral en exceso* [1] [51]. Veamos que consecuencias se derivan de esta relación:

$$\left.\begin{array}{l} |E| = E_\circ e^{\dfrac{\Delta n_I \omega_\circ z}{c_\circ}} \\ \varphi = \omega_\circ \left(t - \dfrac{\Delta n_R z}{c_\circ}\right) \end{array}\right\} \Rightarrow \left.\begin{array}{l} \dfrac{1}{|E|}\dfrac{d|E|}{dz} = \dfrac{\Delta n_I \omega_\circ}{c_\circ} \\ \dfrac{d\varphi}{dz} = -\dfrac{\Delta n_R \omega_\circ}{c_\circ} \end{array}\right\} \Rightarrow \frac{d\varphi}{\dfrac{d|E|}{|E|}} = -\frac{\Delta n_R}{\Delta n_I} = -a \tag{9.20}$$

Esta ecuación nos dice que las variaciones unitarias del módulo del campo eléctrico producen, además, una variación de fase de valor $d\varphi = -a\frac{d|E|}{|E|}$. En la fig. {9.5} se ha representado la variación total de la fase producida por un fotón generado por emisión espontánea, teniendo en cuenta el fenómeno de acoplamiento entre el ruido de amplitud y el de fase. A partir de este resultado, y razonando de la misma manera que en el apartado anterior tenemos:

$$\text{tg}(\Delta\varphi) = \frac{\sin(\theta) - a\cos(\theta)}{\sqrt{P}} \tag{9.21}$$

Como $\varphi \ll 1$ tenemos que $\Delta\varphi \simeq \frac{\sin(\theta) - a\cos(\theta)}{\sqrt{P}}$ con lo que podemos escribir:

$$\begin{aligned} <\Delta\varphi> &= \frac{1}{2\pi}\int_0^{2\pi} \frac{\sin(\theta) - a\cos(\theta)}{\sqrt{P}} d\theta = 0 \\ <\Delta\varphi^2> &= \frac{1}{2\pi}\int_0^{2\pi} \frac{[\sin(\theta) - a\cos(\theta)]^2}{\sqrt{P}} d\theta = \frac{1+a^2}{2P} \end{aligned} \tag{9.22}$$

Por tanto el espectro de emisión del láser vuelve a ser Lorentziano, pero ahora la anchura espectral de emisión $\Delta\omega_l$ viene dada por:

$$\Delta\omega_l = \frac{n_{sp}}{\tau_p 2P} = n_{sp}\frac{\hbar\omega\alpha_E(\alpha_E + \alpha_{sc})c_\circ^2}{4Wn_g^2}(1 + a^2) \tag{9.23}$$

La tasa de emisión espontánea en exceso de un láser a semiconductor es:

$$n_{sp} = \frac{1}{1 - e^{-\frac{F_c - F_v - E_g}{kT}}} \tag{9.24}$$

[1]En inglés "linewidth enhancement factor".

donde F_c y F_v son los cuasiniveles de Fermi en la banda de conducción y de valencia y E_g es la anchura de banda prohibida del semiconductor.

Una cuestión que se plantea de manera natural es por qué el fenómeno de acoplamiento entre el ruido de fase y el ruido de amplitud ocurre sólo en los láseres semiconductores. Después de todo, se trata de una consecuencia de las relaciones de Kramers–Kronig y éstas son válidas para cualquier medio activo. Se puede demostrar, sin embargo, que el coeficiente a es nulo a la longitud de onda correspondiente al máximo de la curva de ganancia del medio *si ésta es simétrica*, cosa que ocurre en los medios activos de los láseres de niveles discretos, pero no en los láseres semiconductores.

9.6. Problemas

Problema 11: Se tiene un láser monomodo semiconductor cuya frecuencia central de emisión es f=192THz. La longitud de la cavidad es L=300μm, el índice de refracción del medio es $n_g = 3,5$ y el coeficiente de pérdidas por unidad de longitud de por scattering Rayleigh es $\alpha_{sc} = 40cm^{-1}$. Teniendo en cuenta que el coeficiente de emisión espontánea en exceso $n_{sp} = 5$ y el coeficiente de acoplamiento entre el ruido de amplitud y el de de fase es $\alpha = 6$, se desea determinar la potencia de emisión necesaria para que la anchura espectral de emisión sea de 50MHz.

Solución:

$$\left.\begin{array}{l} R = \left(\dfrac{n_g - 1}{n_g + 1}\right)^2 \simeq 0,31 \\ \\ \alpha_E = \dfrac{1}{2L}\ln\left(\dfrac{1}{R^2}\right) = 3904m^{-1} \\ \\ \alpha_{sc} = 40cm^{-1} = 4000m^{-1} \end{array}\right\} \Rightarrow \tau_p = \frac{n_g}{(\alpha_E + \alpha_{sc})c_\circ} = 1,5\times 10^{-12}s \Rightarrow \Delta f = \frac{1}{2\pi\tau_p} = 0,106THz$$

$$W = n_{sp}\frac{\hbar\omega\alpha_E(\alpha_E + \alpha_{sc})c_\circ^2}{8\pi\Delta f W n_g^2}(1 + a^2) =$$

$$= \frac{5\times\frac{1,24}{1,55}1,6\times 10^{-19}(3904 + 4000)3904(3\times 10^8)^2}{8\pi\times 50\times 10^3\times 3,5^2}(1 + 36) = 4,27W$$

Capítulo 10

Receptores para comunicaciones ópticas

Resumen: En este capítulo vamos a explicar, en primer lugar, el funcionamiento de los dos fotodetectores que más se emplean en comunicaciones ópticas: el fotodetector PiN y el de avalancha (APD). El primero genera un portador de carga por cada fotón absorbido, mientras que el segundo tiene ganacia interna, es decir, genera, en promedio, un número M$\gg$1 de portadores por fotón detectado. A continuación deduciremos un modelo eléctrico (en el que también se incluirán las fuentes de ruido) de cada uno de los elementos que forman parte del receptor de comunicaciones ópticas: el fotodetector, la resistencia de carga y la primera etapa amplificadora del receptor. A partir de este modelo, calcularemos el cociente entre la potencia de la señal y la potencia del ruido en el receptor para las dos configuraciones más empleadas: el preamplificador de tensión y el de transimpedancia.
Por último, dado que la mayora parte de los sistemas de comunicación por fibra óptica son digitales, relacionaremos, para un caso general, este cociente con la tasa de error binaria máxima admisible del enlace. Veremos, también, cuál ha de ser el valor umbral (o tensión de comparación) óptimo en el circuito de decisión, de manera que se minimice la probabilidad de error del enlace.

10.1. Introducción

En un sistema de comunicaciones ópticas, el receptor tiene como finalidad convertir la señal óptica en una señal eléctrica. El fotodetector (primer elemento del receptor) transforma el flujo de fotones que llegan al extremo de la fibra óptica en una corriente eléctrica (un flujo de electrones) que hay que amplificar para que ésta se distinga claramente del ruido térmico de fondo presente en el receptor.

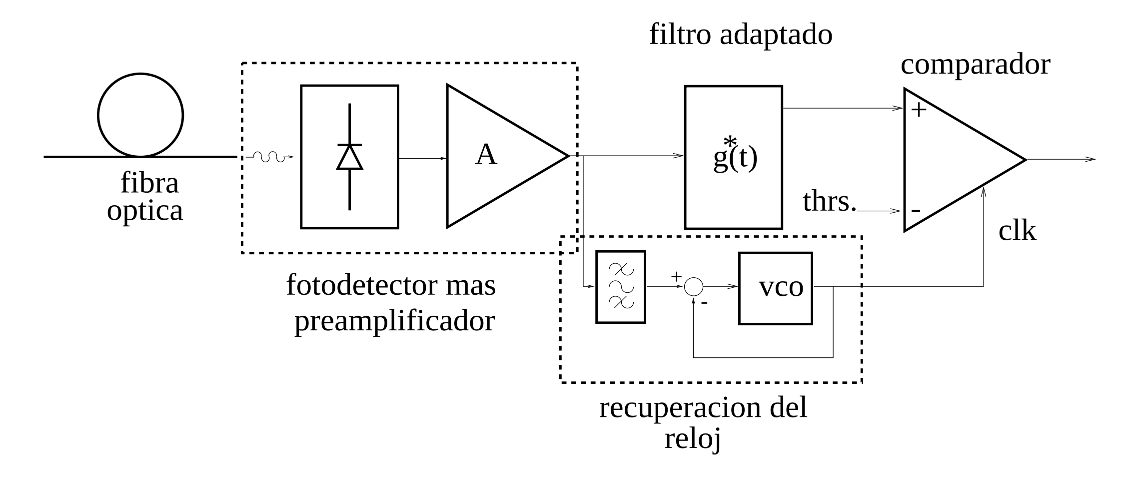


Figura 10.1: Esquema de un receptor de comunicaciones opticas.

En la fig. {10.1} se muestra el diagrama de bloques de un posible receptor de un sistema digital de comunicaciones ópticas. Consta de un fotodetector que a partir de la potencia óptica incidente genera una corriente proporcional a ésta. A continuación hay un bloque amplificador y un filtro adaptado [1] que minimiza la interferencia entre símbolos [83] de forma que la probabilidad de error asociada sea muy baja. La salida del filtro se compara, en el instante de muestreo, con un valor umbral, de manera que si es mayor que éste, se decide que se ha recibido un uno y en caso contrario se decide que se ha recibido un cero. El instante de muestreo lo determina el bloque de recuperación del reloj (PLL, del inglés "Phase locked loop" [48]), que consta de un VCO (un oscilador controlado por tensión [48]) realimentado al que se le da como referencia la salida de un filtro sintonizado a la frecuencia del reloj. En este esquema se supone que se ha utilizado una codificación de los datos de forma que la densidad espectral de potencia a la frecuencia del reloj (o a un múltiplo de ésta) es elevada. Otro esquema muy popular en enlaces digitales consiste en colocar un integrador a la salida del preamplificador y muestrear la salida del integrador al final del tiempo de duración del bit [83]. La interferencia entre símbolos no es mínima, pero se mejora la sensibilidad del receptor ya que la SNR (relación señal a ruido) es mayor.

Como en cualquier otro sistema de comunicaciones, la primera etapa del amplificador es la que limita la sensibilidad del receptor y es a la salida de ésta donde hay que conseguir que el cociente entre la potencia de señal dividido por la potencia del ruido tenga un valor mínimo que garantice una tasa binaria de error inferior a un máximo prefijado. En este capítulo vamos a estudiar los dispositivos fotodetectores y las características de ruido del

[1] Esto es, en muchas ocasiones, muy difícil de llevar a cabo. Unas veces porque no se conoce la función de transferencia del canal y otras porque ésta varía en el tiempo, obligando a implementar un filtro adaptativo.

conjunto formado por el fotodetector y el preamplificador de entrada, ya que ambos definen la sensibilidad del receptor.

10.2. Características de los fotodetectores empleados en comunicaciones ópticas.

El elemento fotodetector se coloca al extremo de la fibra óptica para convertir el pulso luminoso recibido en una pulso eléctrico. La potencia de la señal luminosa recibida puede ser muy débil (del orden del nW), por lo que el detector debe ser capaz de responder ante señales muy débiles (es decir, ser muy sensible). Las propiedades que debe tener un fotodetector para comunicaciones ópticas son las siguientes:

1. Buena sensibilidad a la longitud de onda de trabajo.

2. Una velocidad de respuesta elevada. El convertidor optoelectrónico debe ser capaz de generar una corriente que siga fielmente las rápidas variaciones de los pulsos luminosos.

3. Un bajo nivel de ruido. Puesto que la señal óptica incidente es muy débil, el ruido de origen eléctrico añadido por el fotodetector podría enmascarar la información útil.

4. Dimensiones reducidas, compatibles con el tamaño de la fibra.

5. Fiabilidad y bajo coste.

Los únicos dispositivos que reunen todos estos requisitos son los fotodetectores semiconductores, con ganancia interna o sin ella. El dispositivo sin ganancia interna es el fotodiodo PIN y el que sí la tiene es el fotodetector de avalancha (APD).

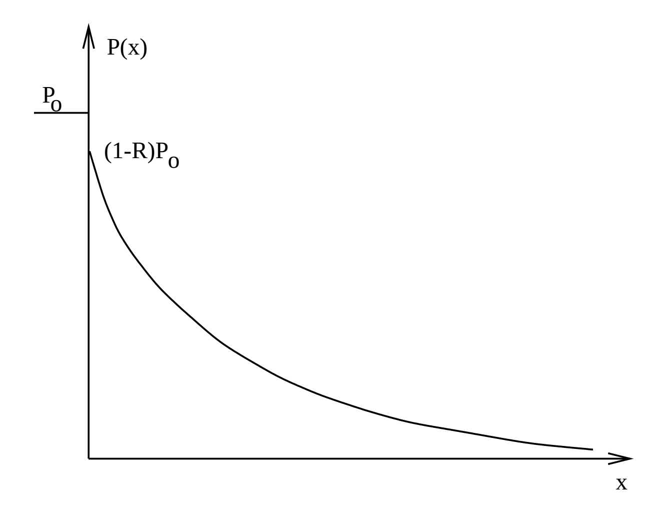


Figura 10.2: Evolución de la potencia óptica en el seno de un semiconductor absorbente.

10.3. Fotoconductores.

Cuando se ilumina un material semiconductor con luz de una frecuencia f tal que la energía de los fotones hf es superior a la anchura de la banda prohibida (E_g), ésta es absorbida creándose pares electrón–hueco [2].

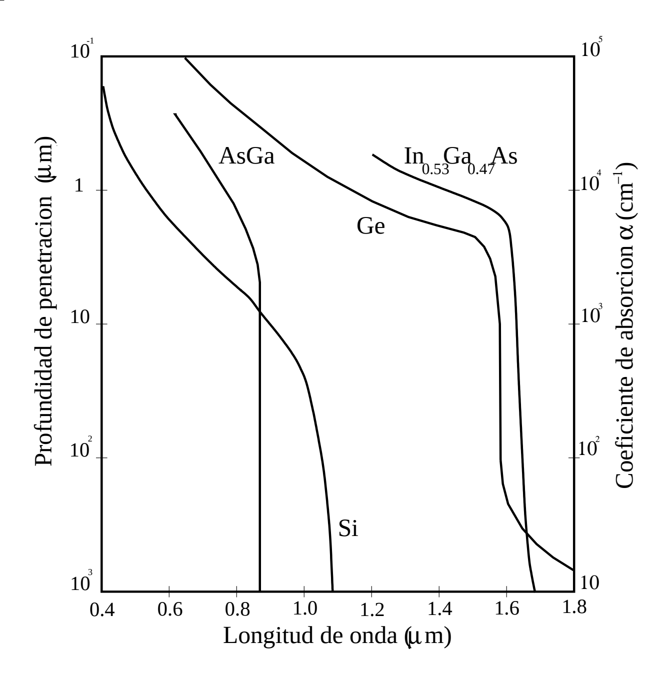


Figura 10.3: Coeficiente de absorción de varios semiconductores.

Si representamos la intensidad luminosa frente a la penetración en el material (ver la fig. {10.2}) [3], observamos que la función presenta una característica exponencial decreciente definida por un coeficiente, denominado coeficiente de absorción del semiconductor y que denotaremos mediante la letra α. La inversa del coeficiente de absorción, $\frac{1}{\alpha}$, tiene dimensiones de longitud y representa la distancia recorrida en promedio por un fotón antes de absorberse. El valor de α para un determinado semiconductor depende de la longitud de onda de la radiación incidente (ver la fig. {10.3}) y de si éste es de transición directa o indirecta. Los semiconductores de transición directa presentan una variación brusca en el coeficiente de absorción a la frecuencia correspondiente a la anchura de la banda prohibida, mientras que en los semiconductores de transición indirecta la variación es más suave [90].

En la fig. {10.4} se muestra un fotoconductor (que no es más que un semiconductor débilmente dopado). Evidentemente la resistencia del semiconductor depende de que esté o no iluminado. Si hay luz presente, la concentración de portadores libres de carga es mayor, por lo que la resistencia será más pequeña. Por tanto, midiendo la resistencia de este dispositivo, podemos determinar la intensidad luminosa incidente.

El problema que presentan los fotoconductores es que una sensibilidad alta exige un

[2]El rendimiento cuántico del proceso de absorción es el cociente entre el número de pares electrón–hueco generados y el número de fotones absorbidos.

[3]La discontinuidad que se observa en $x = 0$ es debida a la fracción de luz que se refleja en la superficie de separación entre el aire y el semiconductor.

tiempo de respuesta grande (dinámica lenta), mientras que un tiempo de respuesta pequeño (dinámica rápida) conlleva una sensibilidad baja. Esto es fácil de comprender si pensamos que cuanto mayor sea el tiempo medio de recombinación entre portadores, mayor será la variación de la resistividad pero su dinámica será más lenta.

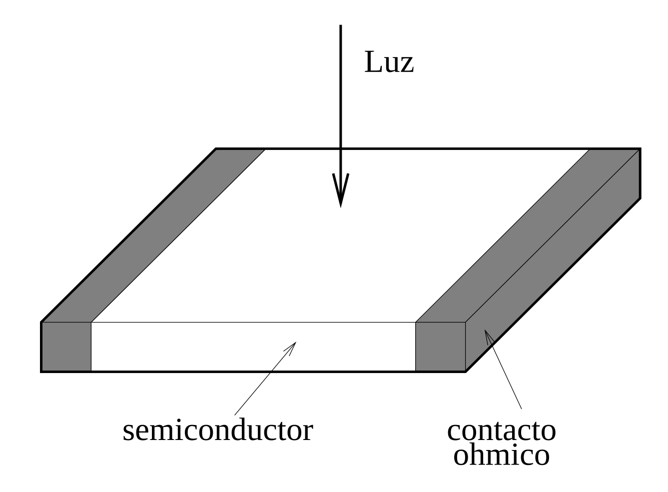


Figura 10.4: Fotoconductor.

Si el fotoconductor estaba iluminado y bruscamente interrumpimos la luz, la concentración de portadores empieza a decrecer exponencialmente con una constante de tiempo igual al tiempo medio de recombinación de los pares electrón–hueco (de igual forma, si el semiconductor estaba a oscuras y bruscamente lo iluminamos la concentración de portadores en desequilibrio empieza a crecer exponencialmente con una constante de tiempo igual al tiempo medio de recombinación τ).

Si el tiempo medio de recombinación de los pares electrón–hueco en desequilibrio es alto, tendremos una sensibilidad apreciable (habrá un cambio grande de resistencia en régimen permanente) pero resulta evidente que su dinámica es inaceptablemente lenta. Si por el contrario el tiempo medio entre recombinaciones es pequeño, su dinámica será rápida, pero su sensibilidad será muy baja.

10.4. El fotodiodo PIN.

Una forma de mejorar las características de sensibilidad y tiempo de respuesta es utilizar un diodo de unión inversamente polarizado. Si la mayor parte de los pares electrón–hueco formados por la absorción de fotones se generan en la región de carga espacial, el fuerte campo eléctrico presente en la zona separa los portadores llevando a cada uno de ellos a la zona donde son mayoritarios, con lo que aparece un portador de carga que atraviesa el dipositivo por cada fotón absorbido (ver la fig. {10.5}).

La corriente que atraviesa el diodo es una corriente inversa de circulación que toma un valor muy pequeño si la unión no está iluminada y será algo mayor si hay luz incidente [77]. En la fig. {10.5} se muestra la curva característica estática del diodo para el caso en que hay luz presente y para el caso en que no la hay.

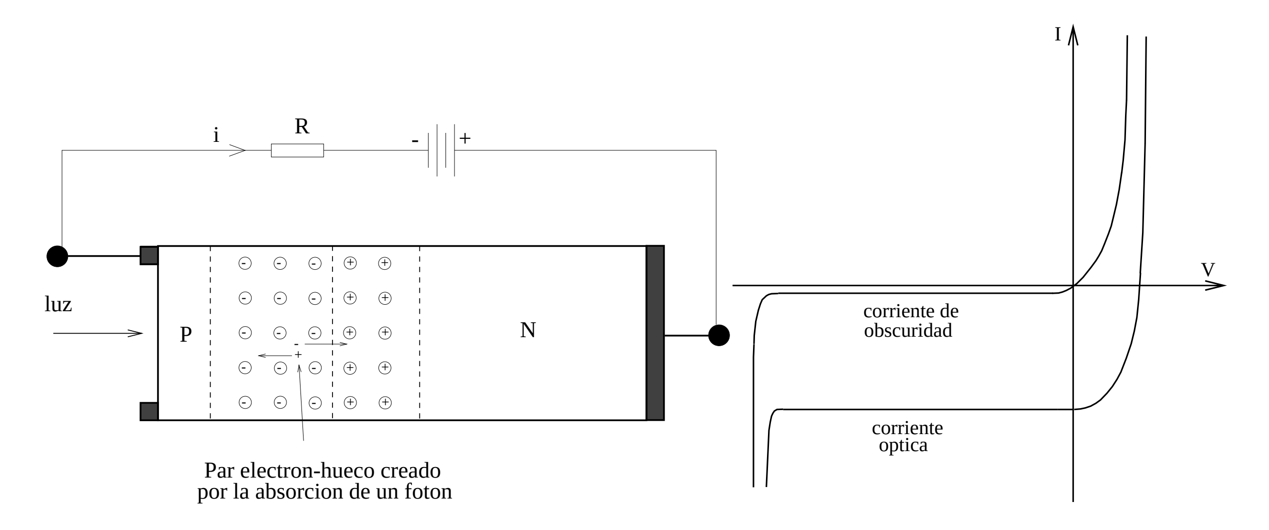


Figura 10.5: Estructura del fotodiodo (izquierda) y curvas características (derecha).

10.4.1. Sensibilidad del dispositivo.

Supongamos que la unión p–n inversamente polarizada (el fotodetector) se ilumina por la zona p. Podemos distinguir tres zonas distintas donde se van a absorber fotones:

- Primero tenemos la zona p, comprendida entre la superficie de entrada y la de comienzo de la región de carga espacial. Aquí la intensidad luminosa $I(x)$ es grande y se generan muchos pares electrón–hueco por unidad de volumen. Para que el el fotón absorbido en esta zona pueda contribuir a la corriente que atraviesa el dispositivo, es necesario que el electrón (que es el portador minoritario) pueda difundirse hasta la región de carga espacial, donde el campo eléctrico lo arrastrará a la zona donde es el portador mayoritario. Esto no será posible si la superficie de entrada del dispositivo tiene una velocidad de recombinación superficial alta [71], ya que entonces el electrón tiende a difundirse en sentido contrario. Por tanto, para aumentar la sensibilidad del fotodetector, es preciso conseguir que esta velocidad de recombinación sea pequeña. En segundo lugar, puesto que queremos que el electrón llegue a la zona de deplexión sin recombinarse en el volumen, la distancia a recorrer deberá ser pequeña en comparación con la longitud de difusión. Para conseguir esto es necesario que el espesor de la zona p sea pequeño frente a la longitud de difusión de los electrones [4] [93]. En resumidas cuentas *la sensibilidad es tanto mayor cuanto menor sea la velocidad de recombinación superficial del plano de entrada y cuanto más estrecha sea la zona p frente a la longitud de difusión de los electrones.*

- En segundo lugar tenemos la región de carga espacial. Los pares electrón–hueco generados aquí se separan y llegan a las zonas volumétricas donde son mayoritarios, contribuyendo así a la fotocorriente que atraviesa el diodo. Interesa que esta región sea lo más ancha posible para que la mayor parte de los fotones se absorban aquí.

- Por último tenemos la zona n. Si la región de deplexión es lo suficientemente ancha podemos despreciar la intensidad luminosa que llega hasta aquí.

[4] El motivo por el que se ilumina por la zona p es que los electrones, que son los portadores minoritarios, tienen coeficientes de difusión más altos que los huecos, por lo que la longitud de difusión es mayor.

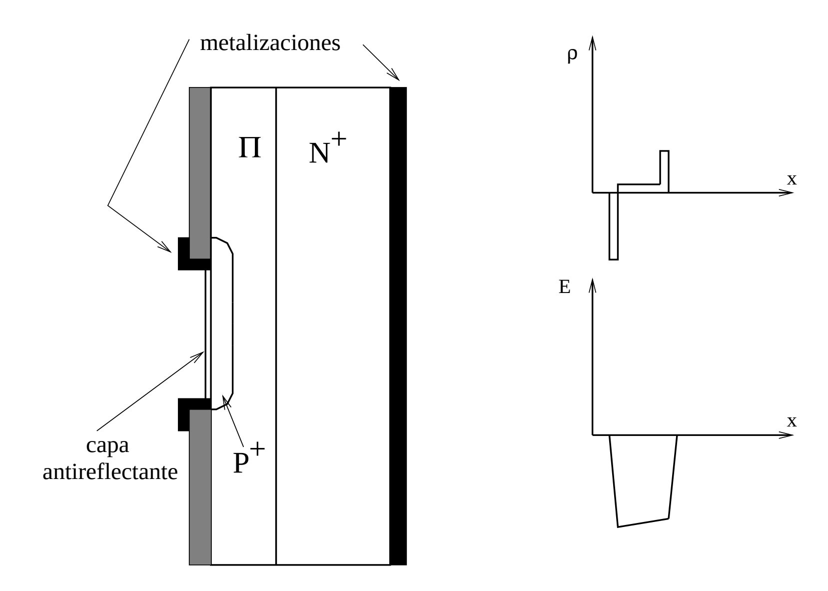


Figura 10.6: Estructura de un fotodiodo pin.

Con objeto de maximizar la sensibilidad del fotodetector se suelen utilizar fotodiodos PIN [77]–[102], es decir, diodos formados por una capa de un semiconductor tipo p fuertemente dopado, una capa casi intrínseca (es decir, en la que el dopado es muy débil [5]) y por último una capa donde hay un fuerte dopado tipo n. En la fig. {10.6} se muestra esta estructura. Las tensiones de polarización inversa necesarias para que la región de carga espacial se extienda por toda la región intrínseca, tal y como se ha representado en la figura, son del orden de la decena de voltios. En la superficie de entrada se intenta que la velocidad de recombinación superficial sea baja [6]. El espesor de la zona p es estrecho con objeto de minimizar la absorción de luz en esta zona, mientras que la región intrínseca es lo suficientemente ancha como para que en ella se absorba la práctica totalidad de la luz.

10.4.2. Tiempo de respuesta.

Entre el momento en que se genera un par electrón–hueco (como consecuencia de la absorción de un fotón) y el momento en que el portador de carga atraviesa la resistencia de carga, contribuyendo así a crear una señal eléctrica detectable y amplificable, transcurre un cierto tiempo, es decir, hay un cierto retraso. Estos retrasos no son iguales para todos los fotones que se han absorbido en un mismo instante, lo que provoca una dispersión de la señal eléctrica con respecto a la óptica o, lo que es lo mismo, una dispersión en el tiempo de respuesta del dispositivo.

- **Tiempo de respuesta interna**. El tiempo de respuesta interno está asociado al tiempo que tardan los portadores en llegar a los contactos metálicos del dispositivo. En realidad existen dos tiempos de respuesta interna, uno asociado a la difusión de los electrones [71] procedentes de la absorción de luz en la zona p (en realidad a la

[5] Cuanto menor es el dopado de un semiconductor mayor es el tiempo medio de recombinación de pares electrón–hueco.

[6] Sin embargo en ésta se suelen depositar una serie de capas de un material dieléctrico con objeto de disminuir el coeficiente de reflexión [68] (que es del orden del 30 % en una interfaz aire–semiconductor). Esta deposición suele aumentar, desgraciadamente, la velocidad de recombinación superficial.

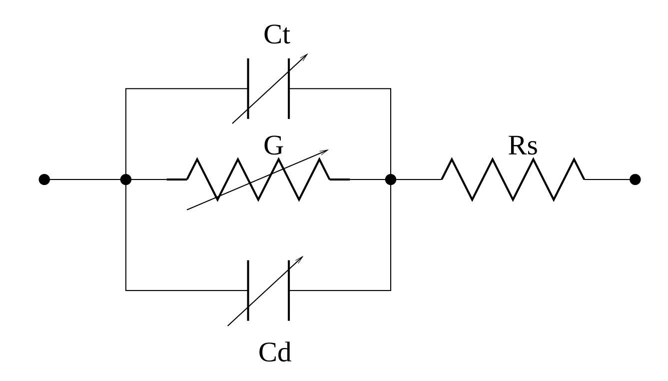


Figura 10.7: Modelo de pequeña señal del fotodetector donde se ha eliminado la fuente de corriente. C_t es la capacidad parásita asociada a la región de carga espacial y C_d la asociada a la variación de carga libre acumulada fuera de la región de carga espacial.

dispersión de estos tiempos) y otro asociado al arrastre de los portadores [71] en la región de carga espacial. El tiempo de respuesta del proceso de difusión es mayor que el del de arrastre. Un valor típico para el primero puede estar en los $100ps$, mientras que el segundo en los $10ps$. En el apéndice 1 se deducen las expresiones analíticas de estos dos tiempos.

- **Tiempo de respuesta externa**. El fotodiodo se polariza inversamente por lo que aparece una capacidad parásita entre sus extremos, denomidada capacidad de transición. La región de carga espacial se comporta como un material dieléctrico que separa dos regiones conductoras (zonas p y n del diodo) y de ahí el comportamiento capacitivo del dispositivo. La expresión del valor de este condensador es:

$$C = \epsilon_o \epsilon_r \frac{A}{w_d} \tag{10.1}$$

donde w_d es la anchura de la región de carga espacial y cuyo valor típico es $200fF$. Normalmente las capacidades parásitas de conexión del dispositivo son más importantes. En cualquier caso, si consideramos el modelo de pequeña señal representado en la fig. {10.7} vemos que la tensión entre los extremos de la resistencia de carga no puede variar más rápidamente que la constante de tiempo del circuito.

En la fig. {10.8} se representa [3] la respuesta temporal de un fotodiodo cuya dinámica está limitada por el tiempo de respuesta interna. La dinámica rápida inicial está asociada al tiempo de tránsito por la región de carga espacial y la lenta al tiempo de difusión de los portadores minoritarios.

10.5. El fotodiodo de avalancha.

Cuando la tensión de polarización inversa aplicada a una unión p–n es suficientemente elevada, los portadores de carga (electrones y huecos) que atraviesan la región de carga

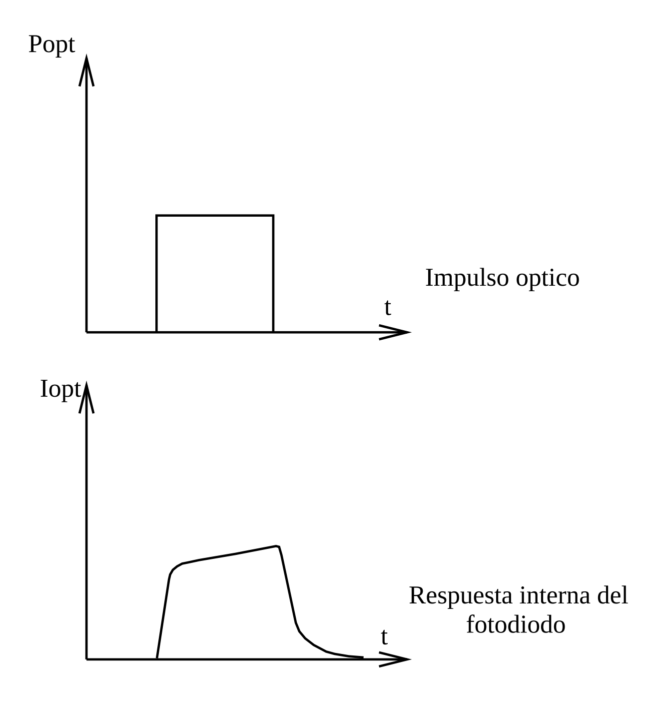


Figura 10.8: Respuesta de un fotodiodo limitado por los tiempos de tránsito y difusión.

espacial pueden adquirir una energía cinética muy grande. Cuando chocan con los núcleos de la red cristalina crean pares electrón–hueco al transferir electrones de la banda de valencia a la banda de conducción. Estos portadores de carga secundarios, generados por choques de los primarios pueden, a su vez, producir otros pares electrón–hueco, produciéndose un fenómeno de avalancha [77]–[64]–[71].

La multiplicación de portadores se puede cuantificar dando dos términos, α y β, denominados coeficientes de ionización, que definen las tasas de creación de pares secundarios a partir de electrones y huecos primarios [64]. Los coeficientes α (relativo a los electrones) y β (relativo a los huecos) dependen del campo eléctrico al que se ven sometido de la siguiente manera [77]:

$$\alpha = A_1 exp(-\frac{A_2}{E}) \tag{10.2}$$

$$\beta = B_1 exp(-\frac{B_2}{E}) \tag{10.3}$$

donde A_1, A_2, B_1 y B_2 son términos que dependen del tipo de semiconductor y de la temperatura.

Los coeficientes de ionización representan la ganancia por unidad de longitud de las densidades de corriente asociadas a los flujos de electrones y huecos debidas al proceso de avalancha, esto es:

$$\frac{dJ_n}{J_n} = \alpha dx \tag{10.4}$$

$$\frac{dJ_p}{J_p} = \beta dx \tag{10.5}$$

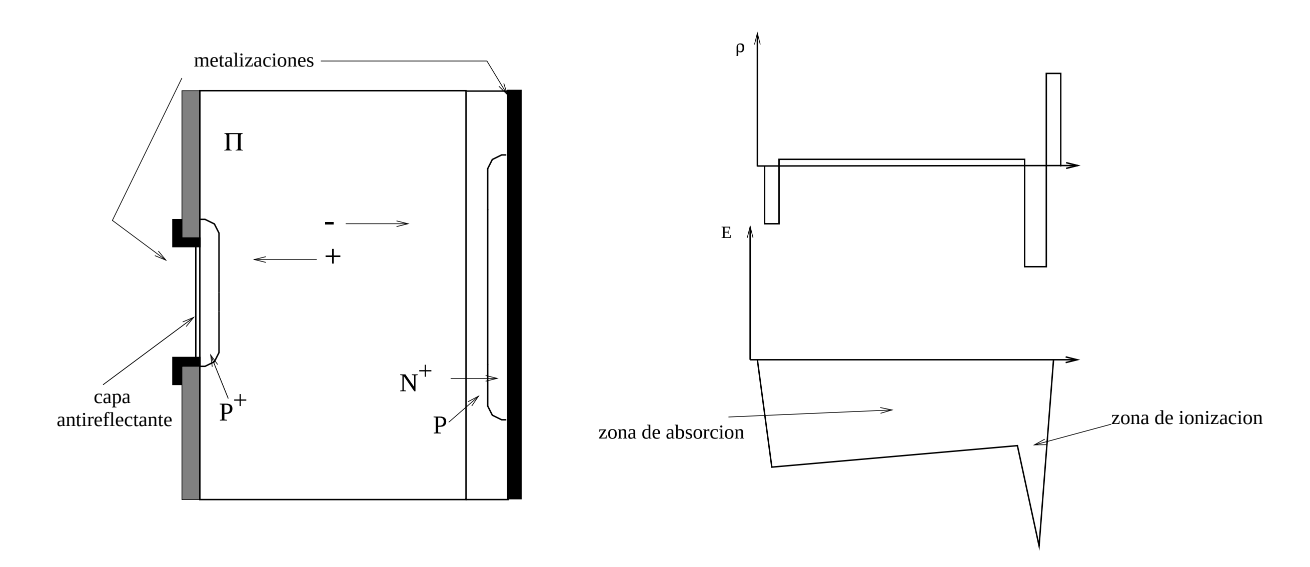


Figura 10.9: Estructura de un fotodiodo de avalancha.

En el fotodiodo de avalancha la ganancia en corriente, el tiempo de respuesta y el ruido añadido a la señal útil dependen fuertemente de los valores que tomen los coeficientes de ionización α y β. Podemos distinguir dos casos límites: uno primero en el que el coeficiente de ionización de los fotoportadores (portadores primarios que, por su mayor movilidad son electrones) es mucho mayor que el de los huecos y uno segundo en el que ambos son del mismo orden de magnitud. Vamos a ver, seguidamente, que para garantizar una dinámica rápida y bajos niveles de ruido la primera situación es la única admisible.

En la fig. {10.9} se muestra la estructura de un fotodiodo de avalancha [102]. Para que la región de carga espacial ocupe toda la región intrínseca y se cree una zona donde se produzca el efecto avalancha, son necesarias tensiones de polarización inversas del orden del centenar de voltios. En la parte intrínseca (zona π) se produce la absorción de fotones, generándose pares electrón–hueco que son separados por el campo eléctrico presente. Los huecos se desplazan hacia la izquierda y los electrones hacia la derecha. Estos electrones (portadores primarios) entran en la unión p^+–n^+ donde, debido al intenso campo eléctrico, provocan, al chocar con la red cristalina, la aparición de pares secundarios produciéndose un efecto avalancha y, consiguientemente, la amplificación de corriente. El coeficiente de amplificación se denota por medio de la letra M y toma, para el caso de APDs de AsGa, valores que son del orden de 100. Si se cumple que $\alpha \gg \beta$ el ruido introducido por el fotodetector no es demasiado grande y además el tiempo de respuesta no empeora demasiado con respecto al fotodiodo PIN.

En la fig. {10.10} se muestra el comportamiento de la región de avalancha para los casos $\alpha \gg \beta$ y $\alpha \simeq \beta$. En el eje horizontal se representa la coordenada longitudinal y en el vertical el tiempo. Puesto que suponemos que los portadores se desplazan a una velocidad de arrastre constante (para una descripción detallada de los fenómenos de transporte en semiconductores remitimos al lector interersado a la bibliografía adjunta [101]–[71]–[90]–[93]), las trayectorias son rectas de pendiente constante (positiva para los electrones y negativa para los huecos, ya que se desplazan con sentidos contrarios). En la primera figura se observa que sólo los electrones generan pares electrones–hueco secundarios, ya que su coeficiente de ionización es mucho mayor que el de los huecos. Esto lleva a que el retraso máximo entre

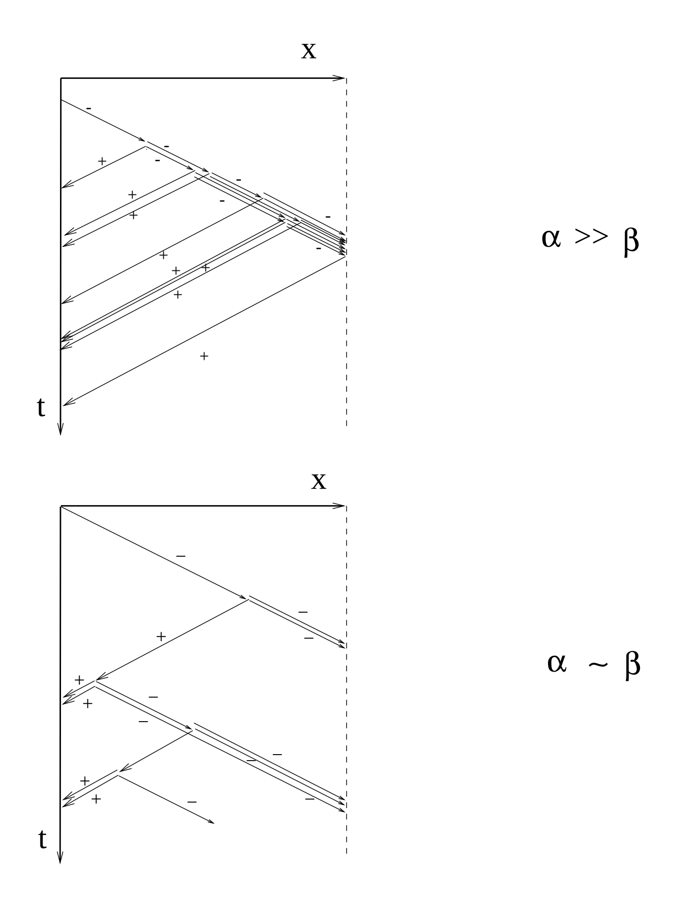


Figura 10.10: Comportamiento de la región de avalancha.

el instante en que entra el portador primario en la región de avalancha y el momento en que deja de haber portadores secundarios provocados por él sea, aproximadamente, la suma del tiempo de tránsito del electrón más el tiempo de tránsito del hueco. Además el ruido introducido por el efecto de avalancha está determinado sólo por el carácter aleatorio de la generación de pares secundarios.

En la segunda figura se observa que no podemos acotar de una forma sencilla el retraso máximo y que el número de portadores secundarios generados por un portador primario es mucho más variable que en el caso anterior. Por este motivo las características de ruido y tiempo de respuesta son mediocres.

10.6. Fuentes de ruido en un receptor.

La fig. {10.1} representa el diagrama de bloques de un receptor de comunicaciones ópticas. Como ya hemos visto, el fotodetector convierte la potencia óptica recibida en una corriente eléctrica. En el siguiente bloque del receptor, el amplificador, se convierte esta corriente en una tensión que debe ser lo suficientemente intensa como para poder ser tratada en el bloque demodulador. Como en cualquier otro sistema de comunicaciones, la primera etapa del amplificador es la que limita la sensibilidad del receptor. En este punto hay que garantizar una relación potencia de señal–potencia de ruido mínima que garantice la calidad de la transimisión. En la fig. {10.11} se representa la etapa de entrada del receptor, así como

su modelo de pequeña señal [7].

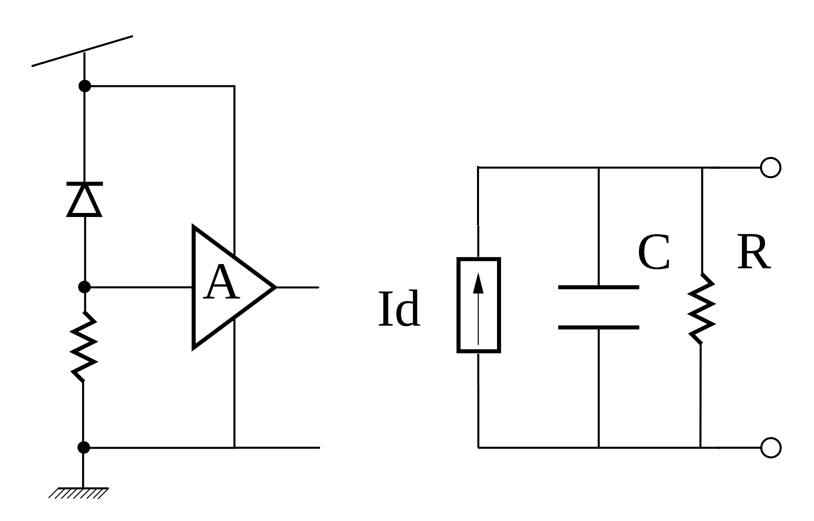


Figura 10.11: Conjunto formado por el fotodetector y el preamplificador.

Existen varios fenómenos físicos que introducen ruidos (señales de carácter aleatorio y media cero) que no están incluidas en este modelo. Para analizar la sensibilidad del receptor, vamos a estudiar las características de las distintas fuentes de ruido para después modelarlas eléctricamente [8] e incluirlas en el circuito de pequeña señal.

1. En primer lugar tenemos el llamado **ruido cuántico asociado a la propia señal**. Es el único término de ruido que es imposible eliminar, ni tan siquiera en un experimento idealizado ya que está asociado a la naturaleza cuántica de la propia señal útil. Consideremos que se detecta una pulso luminoso de potencia P y duración τ. En el proceso de detección, el pulso luminoso se comporta como un conjunto de corpúsculos, llamados fotones, cuyo número no es determinista sino aleatorio. Este carácter aleatorio es el responsable del ruido cuántico asociado a la señal recibida.

2. **Ruido térmico de la resistencia de carga.** Está provocado por el movimiento aleatorio de los portadores de carga (electrones) en la resistencia de carga R_L. Este movimiento, de origen térmico, provoca fluctuaciones de tensión en los extremos de la resistencia de carga. Teóricamente podría eliminarse enfriando ésta a la temperatura de 0K.

3. **Ruido térmico del amplificador.** Cualquier amplificador degrada el cociente $\frac{S}{N}$ de la señal que está tratando, debido al ruido añadido por los elementos activos con los que se ha construido.

4. **Ruido cuántico asociado a la corriente de obscuridad.** Una corriente eléctrica está provocada por el paso de portadores discretos de carga (electrones y huecos en el caso de un semiconductor) a través del dispositivo. El carácter corpuscular de éstos

[7] En la mayora parte de las ocasiones la señal luminosa que llega al receptor es lo suficientemente débil como para que el modelo linealizado (de pequeña señal) de la etapa de entrada describa correctamente su comportamiento.

[8] El modelo eléctrico de una fuente de ruido consiste en una fuente de corriente de media cero y valor cuadrático medio conocido. El valor cuadrárico medio representa la potencia del ruido.

provoca fluctuaciones de la corriente alrededor de su valor medio. En el caso de un fotodiodo, sabemos que existe una cierta corriente que lo atraviesa en ausencia de luz (corriente inversa de saturación, también llamada corriente de obscuridad) que es tanto más importante cuanto menor sea la anchura de la banda prohibida del semiconductor del que está hecho el fotodiodo. Las fluctuaciones de la corriente de obscuridad alrededor de su valor medio introducen un término de ruido que puede ser importante si la corriente inversa de saturación es (relativamente) grande [9]. Sin embargo no vamos a considerar esta fuente de ruido en nuestro análisis.

5. Si el fotodetector que se emplea es un fotodiodo de avalancha aparece un factor de ruido adicional: **el ruido de avalancha**, provocado por el hecho de que el número de portadores secundarios generados a partir de uno primario es variable, por lo que se introduce un término de ruido en el proceso de amplificación interna.

En los próximos apartados vamos a estudiar cada uno de estos término, a modelarlos e incluirlos en el modelo de pequeña señal del receptor [41]-[57].

10.6.1. Ruido cuántico asociado a la señal.

Proceso de detección.

El proceso de detección de la luz impone sobre ésta un comportamiento corpuscular [10]. La tasa media de llegada fotones al detector se puede calcular fácilmente. Si P es la potencia luminosa, el número de fotones medios detectados por unidad de tiempo es:

$$r = \frac{P}{hf} \tag{10.6}$$

Si admitimos que los tiempos de llegada de los fotones no están temporalmente correlados, se puede demostrar que el proceso estadístico definido por las probabilidades discretas $p_n(t)$ de recibir n fotones al cabo de t segundos es poissoniano. Por consiguiente [52]–[69]–[89]:

$$p_n(t) = \frac{(rt)^n}{n!} e^{-rt} \tag{10.7}$$

Si particularizamos para $t = \tau$ (tiempo de definición de un bit) y tenemos en cuenta que el producto $r\tau$ es el número medio (esperado) de fotones por cada pulso luminoso, entonces resulta que el número de fotones que vienen en un pulso luminoso se describe mediante una variable aleatoria discreta definida por las probabilidades:

$$p_n = \frac{\overline{n}^n}{n!} e^{-\overline{n}} \tag{10.8}$$

donde $\overline{n} = \frac{P\tau}{hf}$.

[9] Como es el caso de los fotodetectores de Germanio, ya que la banda prohibida de este semiconductor es relativamente estrecha.

[10] Los quantos, entes físicos descritos por la mecánica cuántica, manifiestan una naturaleza ondulatoria o corpuscular dependiendo del experimento en que intervienen.

La señal útil transmitida está asociada al valor medio de la variable aleatoria, mientras que el ruido cuántico está asociado a las fluctuaciones que sufre el número de fotones detectados en torno al valor esperado. Puesto que conocemos las expresiones de las probabilidades discretas, podemos calcular fácilmente la desviación cuadrática media de la variable aleatoria. Haciendo las operaciones correspondientes, resulta:

$$\overline{\Delta n^2} = \overline{n} \tag{10.9}$$

Conociendo el valor esperado del número de fotones por pulso y la desviación cuadrática media podemos obtener fácilmente el modelo "ruidoso" de la señal eléctrica generada por el detector. La señal útil está asociada a la media de la variable aleatoria. La corriente media que atraviesa el fotodetector sería igual al producto del número medio de pares electrón–hueco generados por la absorción de fotones, multiplicados por el rendimiento cuántico del fotodetector y por la carga de un portador y dividido por el tiempon de duración del pulso luminoso, es decir:

$$\overline{i} = \frac{q\eta\overline{n}}{\tau} = \frac{q\eta P}{hf} \tag{10.10}$$

La potencia de la señal eléctrica útil es igual al cuadrado de la expresión anterior. La potencia del ruido eléctrico la podemos calcular fácilmente (suponiendo que η es una constante y no una variable aleatoria):

$$\overline{\Delta i^2} = \frac{q^2\eta^2\overline{\Delta n^2}}{\tau^2} = \frac{q^2\eta P}{hf\tau} \tag{10.11}$$

Si ahora elevamos al cuadrado (10.10) y dividimos el resultados por (10.11) obtenemos la expresión del cociente potencia útil de la señal dividida por la potencia del ruido cuántico asociado a ésta:

$$\frac{S}{N} = \frac{\overline{i}^2}{\overline{\Delta i^2}} = \frac{P\tau\eta}{hf} \tag{10.12}$$

Vemos, pues, que el fotodetector necesariamente degrada la relación señal a ruido, a menos que tenga características ideales ($\eta = 1$). El ruido cuántico existe en todos los sistemas de comunicaciones, aunque sólo en comunicaciones ópticas es preciso tenerlo en cuenta. Esto se puede justificar fácilmente a partir de (10.12). En el dominio radioeléctrico las longitudes de onda son lo suficientemente largas (y, por consiguiente, las frecuencias lo suficientemente bajas) como para que el producto hf sea despreciable. A estas frecuencias tan bajas el carácter corpuscular de la energía detectada no es observable [11]. En comunicaciones ópticas la energía de un cuanto de luz (de un fotón) es apreciable frente a la del pulso luminoso [12], por lo que este carácter corpuscular se pone de manifiesto, apareciendo un término de ruido adicional, el *ruido cuántico.*

[11] Esto es así porque en un pulso siempre hay un número enorme de cuantos.

[12] Siempre y cuando la energía del pulso luminoso, es decir, el producto $P\tau$, sea suficientemente pequeño.

Modelo eléctrico.

Para modelar el proceso de detección, se consideran dos fuentes de corriente en paralelo, una de las cuales representa la señal útil mientras que la otra representa el ruido cuántico asociado a ésta. El valor medio de la fuente de corriente asociada al ruido es 0, mientras que el de la señal útil es [13]:

$$\begin{aligned} \bar{i} &= RP \\ R &:= \frac{q\eta}{hf} \end{aligned} \qquad (10.13)$$

La otra fuente representa las fluctuaciones alrededor de este valor medio. Teniendo en cuenta (10.10) y (10.11) llegamos a [41]:

$$\overline{\Delta i^2} = \frac{q\bar{i}}{\tau} = 2q\bar{i}\Delta f \qquad (10.14)$$

donde Δf es el ancho de banda de la envolvente de la señal óptica. Normalmente el ancho de banda del amplificador de entrada es igual al de la señal por lo que, a partir de ahora, consideraremos que Δf es el ancho de banda eléctrico del amplificador.

10.6.2. Ruido térmico en la resistencia de carga.

Los portadores de carga de una resistencia (electrones) están sometidos a un movimiento de agitación térmica continuo. Si colocamos dos resistencias en paralelo, tendremos que la energía máxima que, por unidad de tiempo y de intervalo espectral, puede ceder una a la otra viene dada por [14]:

$$\widetilde{P} = \frac{hf\,df}{e^{\frac{hf}{kT}} - 1} \qquad (10.15)$$

Si tomamos como temperatura 290K tenemos que la potencia total máxima [15] que se podría ceder es:

$$P_{tot} = \int_0^\infty \frac{hf}{e^{\frac{hf}{kT}} - 1} df = 40nW \qquad (10.16)$$

Esta número es muy grande y, si realmente éste fuese el valor de la potencia debida al ruido térmico, la señal útil estaría, en la mayor parte de los casos, totalmente enmascarada.

[13] Esta expresión se deduce de modo inmediato de (10.10) y de la definición de R (que recibe el nombre de responsividad del fotodetector).

[14] Esta expresión es una consecuencia inmediata de la ley de Planck acerca de la radiación de un cuerpo negro.

[15] La potencia transferida sería máxima siempre que hubiera adaptación de impedancias, es decir, que la segunda resistencia fuese de igual valor que la primera.

Sin embargo el ancho de banda del amplificador es limitado, por lo que en la práctica la potencia debida al ruido térmico es mucho menor ya que se filtran las componentes de más alta frecuencia. Para calcularla bastaría con calcular la integral anterior sustituyendo ∞ por la frecuencia de corte del amplificador B.

En realidad el cálculo de la potencia del ruido térmico se puede simplificar mucho. A temperatura ambiente $kT = 0{,}026eV$ lo que se corresponde con una frecuencia $f \simeq 3THz$, es decir un valor muy grande en comparación con la de corte del amplificador [16]. Esto nos lleva a que $hf \ll kT$, con lo que podemos aproximar la densidad espectral de potencia del ruido térmico mediante la conocida expresión [17]:

$$\widetilde{P} = kT\Delta f \tag{10.17}$$

Una vez conocida esta expresión es inmediato determinar el modelo de ruido de la resistencia de carga. Este consiste en una fuente de corriente de media 0 y valor cuadrático medio $\overline{\Delta i^2} = \frac{4kT}{R}\Delta f$ en paralelo con la resistencia, ya que de esta manera si conectáramos una resistencia de carga igual (adaptación de impedancias), se disiparía una potencia de valor $kT\Delta f$.

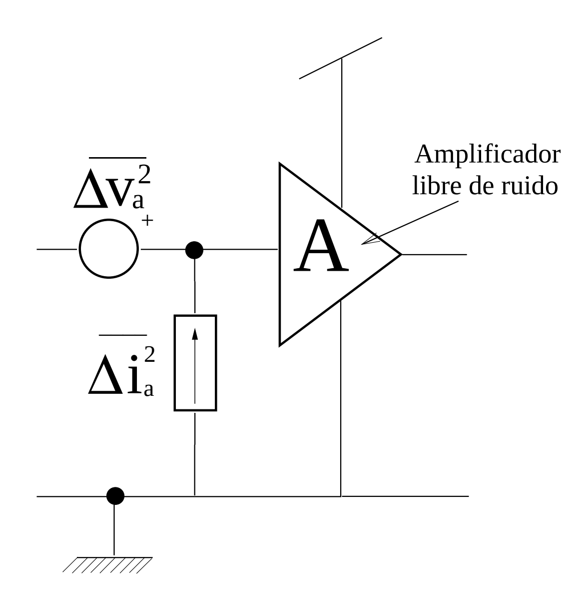


Figura 10.12: Modelo de ruido del preamplificador.

10.6.3. Ruido del preamplificador.

La función del preamplificador es la de producir una señal lo suficientemente intensa como para que la puedan tratar las etapas posteriores del receptor. La amplificación introduce un ruido adicional de origen complejo [18] que podemos modelar por medio de dos

[16] Esta frecuencia puede ser, como mucho, del orden de algunas decenas de Ghz.

[17] Vemos que en esta aproximación, válida para el dominio radioeléctrico, el ruido térmico se puede considerar espectralmente blanco.

[18] Intervienen fuentes de ruido cuántico y térmico, entre otras.

fuentes que *vamos a suponer independientes entre sí* [19], una de intensidad y otra de tensión, colocadas a la entrada del amplificador de la manera indicada en la fig. {10.12}. Las medias cuadráticas de estas dos fuentes de ruido dependen del tipo de dispositivo que se emplee [41] (transistor bipolar de Si, JFET de Si o MESFET de AsGa) y de las condiciones de polarización. Normalmente las condiciones de polarización óptimas en cuanto a niveles de ruido no coinciden con una ganancia máxima. Los valores típicos de los parámetros de ruido de un transistor bipolar de Si polarizado en condiciones óptimas son $\sqrt{\overline{\Delta i_a^2}} = \frac{2pA}{\sqrt{Hz}}$ y $\sqrt{\overline{\Delta V_a^2}} = \frac{2nV}{\sqrt{Hz}}$. Los de un transistor J-FET de Si son $\sqrt{\overline{\Delta i_a^2}} = \frac{10fA}{\sqrt{Hz}}$ y $\sqrt{\overline{\Delta V_a^2}} = \frac{4nV}{\sqrt{Hz}}$. Por último, los de un MES–FET de AsGa son $\sqrt{\overline{\Delta i_a^2}} = \frac{100fA}{\sqrt{Hz}}$ y $\sqrt{\overline{\Delta V_a^2}} = \frac{1nV}{\sqrt{Hz}}$. Este tipo de transistores tienen capacidades de entrada muy bajas, del orden de $0{,}5pF$.

10.6.4. Ruido de amplificación interna (ruido de avalancha).

El efecto de avalancha hace que por cada fotón absorbido se generen, como media, $\overline{M}$ portadores de carga que atraviesan el dispositivo. Sin embargo, el número real, en un caso particular, viene representado por una variable aleatoria que tiene una determinada desviación cuadrática media. Estas fluctuaciones provocan una degradación del cociente potencia señal–potencia ruido. Esta degradación viene determinada por el siguiente coeficiente [102]:

$$F := \frac{\overline{M^2}}{\overline{M}^2} = 1 + \frac{\overline{\Delta M^2}}{\overline{M}^2} \tag{10.18}$$

Los valores típicos de ganancia en corriente y factor de ruido de un fotodiodo APD de Si (empleado en la primera ventana de transmisión) son $\overline{M} = 100$ y $F = 6$, mientras que los de un APD de Ge (empleados en la seguda y tercera ventana de transmisión) son $\overline{M} = 20$ y $F = 10$.

10.7. Cálculo de la relación potencia señal–potencia ruido.

Ya estamos en condiciones de incluir los términos debidos a las fuentes de ruido en el modelo de la fig. {10.11}. El circuito resultante se muestra en la fig. {10.13}.

La resistencia R (resistencia equivalente al conjunto formado por la resistencia de carga del fotodetector y la resistencia de entrada del preamplificador) y la capacidad C (suma de la capacidad parásita del fotodetector más la capacidad de de entrada del preamplificador) definen un filtro paso bajo de frecuencia de corte $\frac{1}{2\pi RC}$ que puede distorsionar la señal recibida. En principio sería deseable que esta frecuencia fuese superior al ancho de banda eléctrico de la señal que se transmite. Para conseguir esto se procura minimizar durante el proceso de fabricación el valor de las capacidades parásitas. Una manera obvia de conseguir que el preamplificador tenga un ancho de banda grande es tomar una resistencia de carga

[19] Esto no es cierto a altas frecuencias. Sin embargo, con objeto de facilitar los cálculos, vamos a admimitir esta simplificación.

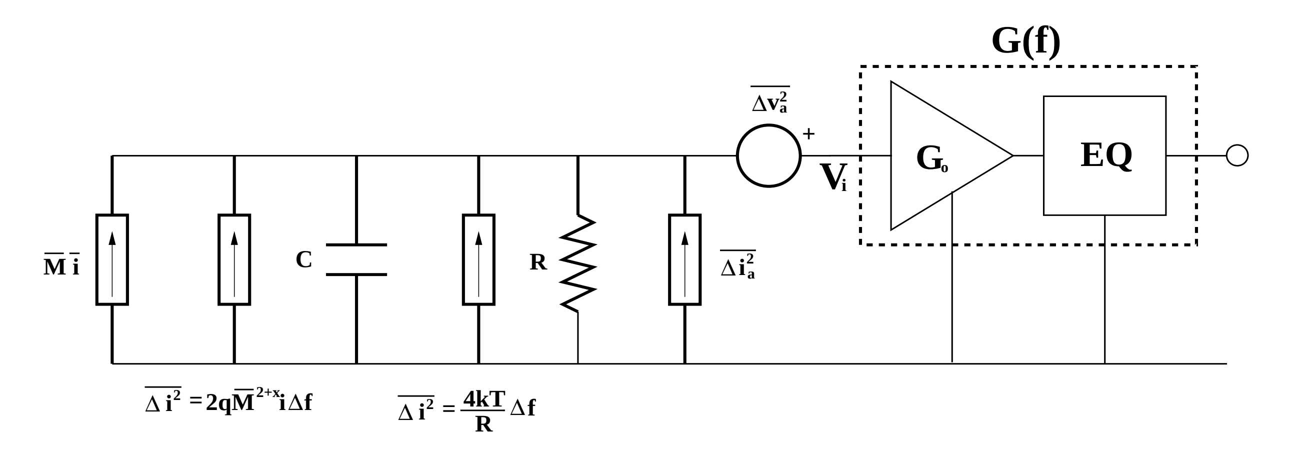


Figura 10.13: Modelo de ruido del receptor completo.

pequeña. El problema es que entonces el ruido térmico introducido por esta resistencia se hace también grande.

Hay casos (transmisión de grandes volúmenes de información a grandes distancias [20]) en las que resulta inevitable que la resistencia R y el condensador C filtren la señal que se recibe. Para corregir esto es preciso añadir un ecualizador que amplifique más las componentes de alta frecuencia que se han filtrado durante la detección. Este ecualizador se ha incluido en la fig. {10.13}.

Si hacemos caso omiso de los términos de ruido, tendremos que la tensión a la entrada del preamplificador es [21]:

$$V_i = \overline{M}\overline{i}Z_{eq} = \overline{M}\overline{i}\frac{1}{\frac{1}{R} + j2\pi fC} = \frac{R\overline{M}\overline{i}}{1 + j2\pi fRC} \tag{10.19}$$

La red de ecualización debe estar diseñada para corregir este filtrado, de manera que la señal a la salida del preamplificador sea:

$$V_\circ = V_i G(f) = \frac{R\overline{M}\overline{i}}{1 + j2\pi fRC} G_\circ(1 + j2\pi RC) = G_\circ R\overline{M}\overline{i} \tag{10.20}$$

Si ahora tenemos en cuenta sólo las fuentes de ruido, podemos calcular la desviación típica [22] de las fluctuaciones alrededor de la tensión media obtenida anteriormente. Supondremos que las fuentes de ruido son independientes, con lo que podemos sumar las desviaciones cuadráticas medias producidas por cada una de las fuentes de ruido. Vamos a agrupar todas las fuentes de corriente de ruido en una sola, cuya media cuadrática es la suma de las medias cuadráticas de cada una de las componentes (consideramos que las variables

[20] En estos casos es inevitable tomar una resistencia de carga grande ya que si no se hiciera esto el ruido térmico introducido por R enmascararía la señal útil.

[21] Estamos considerando el caso más general en el que el fotodetector sea un fotodiodo de avalancha. Para particularizar para un fotodiodo PIN basta con tomar $\overline{M} = 1$.

[22] La desviación típica de una variable aleatoria es la raíz cuadrada de la desviación cuadrática media.

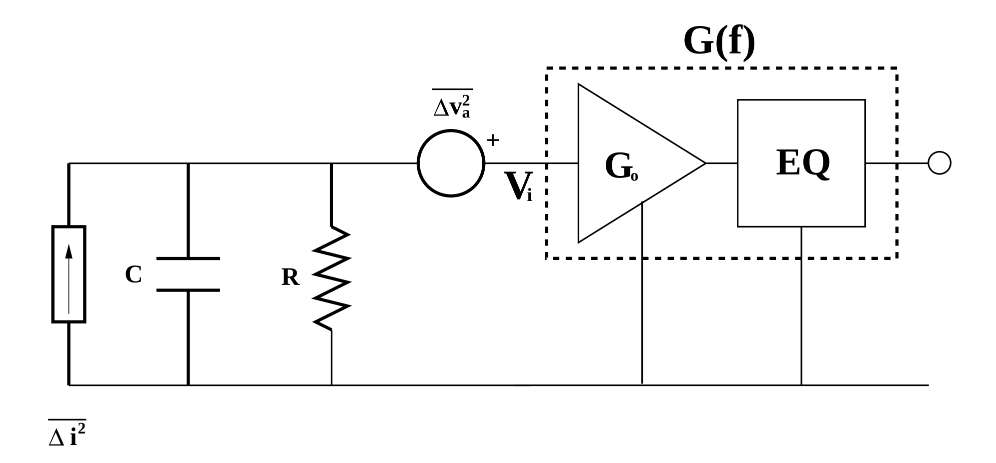


Figura 10.14: Modelo equivalente de ruido del receptor.

aleatorias son independientes). Por tanto el modelo de ruido equivalente es el mostado en la fig. {10.14}.

Teniendo en cuenta que tenemos dos fuentes independientes, una de tensión y otra de intensidad, podemos deducir, aplicando el principio de superposición, la expresión de las fluctuaciones cuadráticas medias de tensión (por unidad de anchura espectral) a la entrada del amplificador–ecualizador:

$$\overline{\Delta V_i^2} = \overline{\Delta i^2}\frac{R^2}{1+4\pi^2 f^2 R^2 C^2}\Delta f + \overline{\Delta V_a^2}\Delta f \tag{10.21}$$

A partir de esta expresión calculamos las fluctuaciones cuadráticas medias *totales* a la salida:

$$\overline{\Delta V_{\circ}^2} = \int_0^B \frac{R^2\overline{\Delta i^2}}{1+4\pi^2 f^2 R^2 C^2}|G(f)|^2 df + \int_0^B \overline{\Delta V_a^2}|G(f)|^2 df \tag{10.22}$$

Desarrollando esta expresión [23] se deduce la siguiente expresión para la desviación típica de la tensión de salida:

$$\sqrt{\overline{\Delta V_o^2}} = G_o R\sqrt{\overline{\Delta i^2} + \overline{\Delta V_a^2}\left[\frac{1}{R^2} + \frac{4\pi^2}{3}(\Delta f)^2 C^2\right]}\sqrt{\Delta f} \tag{10.23}$$

Teniendo ahora en cuenta las distintas componentes que intervienen en la fuente de ruido $\overline{\Delta i^2}$, se llega a que el cociente [41] entre el valor medio de la tensión de salida dividido por la desviación típica de sus fluctuaciones, cociente que denotamos mediante la letra K, tiene la siguiente expresión [24]:

[23] Aconsejamos al lector que haga el desarrollo.

[24] De nuevo hemos considerado el caso más general en el que el fotodetector es un fotodiodo de avalancha. Para contemplar el caso de un fotodiodo PIN basta con hacer $\overline{M} = 1$ y $F = 1$.

$$K := \frac{V_o}{\sqrt{\overline{\Delta V_o}}} = \frac{\bar{i}}{\sqrt{2qF\bar{i} + \frac{4kT}{R\overline{M}^2} + \frac{\overline{\Delta I_a^2}}{\overline{M}^2} + \frac{\overline{\Delta V_a^2}}{\overline{M}^2}\left[\frac{1}{R^2} + \frac{4\pi^2}{3}(\Delta f)^2 C^2\right]}\sqrt{\Delta f}} \tag{10.24}$$

De esta expresión podemos deducir algunas conclusiones de carácter general:

1. El cociente K crece a medida que aumenta el coeficiente de avalancha $\overline{M}$, hasta que llega un momento en el que el ruido de avalancha, que está multiplicado por el factor F, se hace dominante.

2. Un aumento en la resistencia de carga R aumenta el cociente K, siempre que alguno de los términos en los que interviene sea el dominante. Sin embargo, utilizar resistencias de carga de valor elevado tiene un inconveniente obvio: es preciso añadir un circuito de ecualización, lo que limita el rango dinámico de funcionamiento del receptor. El rango dinámico indica la capacidad de manejar simultáneamente señales fuertes y débiles. En un esquema como el descrito hasta ahora, el rango dinámico es pequeño ya que para conseguir una sensibilidad grande *en toda la banda* de paso es preciso amplificar mucho en baja frecuencia (antes del polo definido por R y C). En el caso de que la señal de entrada tome un valor elevado (si el rango dinámico es alto, habrá veces en las que la señal de entrada será débil y otras en las que será intensa), las componentes de baja frecuencia pueden saturar el preamplificador provocando distorsión.

3. A altas frecuencias, el término capacitivo del denominador puede hacerse dominante, ya que crece con el cuadrado de la frecuencia. Este es un motivo adicional para intentar minimizar C.

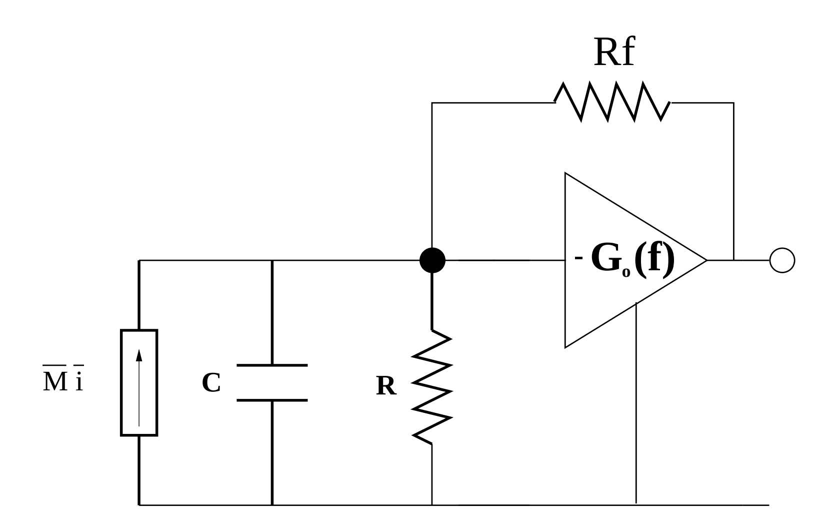


Figura 10.15: Preamplificador de transimpedancia.

La configuración del preamplificador que hemos estudiado hasta ahora (la de un amplificador de tensión en el que la resistencia de carga sirve como convertidor intensidad–tensión), no es la que se utiliza más habitualmente. Con mayor frecuencia se emplea el circuito representado en la fig. {10.15}, el amplificador de transimpedancia [86].

Si calculamos la función de transferencia de este circuito, llegamos a la siguiente expresión:

$$V = -\frac{\dfrac{R_f\overline{M}}{1+\frac{1}{G_\circ}+\frac{R_f}{G_\circ R}}}{1+j\dfrac{2\pi f C R_f}{1+\frac{R_f}{R}+G_\circ}}\bar{i} \tag{10.25}$$

Si la ganancia del amplificador verifica $G_\circ \gg (1+\frac{R}{R_f})$ la expresión anterior se simplifica, quedando:

$$V \simeq \frac{-R_f\overline{M}\bar{i}}{(1+\frac{j2\pi f R_f C}{G_\circ})} \tag{10.26}$$

Por lo que si además se toma una ganancia lo suficientemente alta como para que $G_\circ \gg 2\pi C R_f \Delta f$, resulta que no es preciso ecualizar ya que

$$V \simeq -R_f\overline{M}\bar{i} \tag{10.27}$$

La ventaja de estos preamplificadores consiste en que, al no ser necesario ecualizar su respuesta, tienen un rango dinámico de funcionamiento mayor que el del amplificador de tensión. Como desventaja, sus características de ruido son peores y los requerimientos de diseño del amplificador son mucho más estrictos (la ganancia debe ser muy alta en todo el ancho de banda). Se puede demostrar que la expresión (10.24) sigue siendo válida para los preamplificadores de transimpedancia, a condición de sustituir la resistencia R por el paralelo de las resistencia R y R_f, lo que hace que aumente el ruido térmico.

10.8. Tasa de error binario máxima admisible.

En la fig. {10.16} se ha representado las funciones de densidad de probabilidad (al haber ruidos aleatorios superpuestos las señales no son deterministas) asociadas a los siguientes sucesos: recibir un símbolo correspondiente a un cero (denotado en la figura como "off") y recibir un símbolo correspondiente a uno (denotado en la figura como "on"). También se ha representado el valor umbral de decisión. Si la tensión medida en el bloque de decisión (comparador) es mayor que V_{th} se asume que se ha recibido recibido un uno lógico y en caso contrario un cero lógico. Debido al ruido presente en el receptor, puede ocurrir que se haya recibido un cero y el ruido superpuesto a la señal haga que la tensión sea superior a la umbral. En este caso decidiremos, erróneamente, que se hemos recibido un uno. También puede ocurrir que llegue un uno y por culpa del ruido, la tensión a la entrada del comparador sea inferior a V_{th}. También en este caso concluiremos que hemos recibido el símbolo equivocado (un cero en este caso). Vamos a suponer que los ruidos superpuestos a la señales útiles (representadas por los valores medios μ_{on} y μ_{off}) son gaussianos; denotaremos la probabilidad de recibir un "1" por medio de la letra a y la de recibir un "0" por medio de $1-a$ y las probabilidades de decidir, erróneamente, que ha llegado un "1" por medio de $P_\circ$ y la

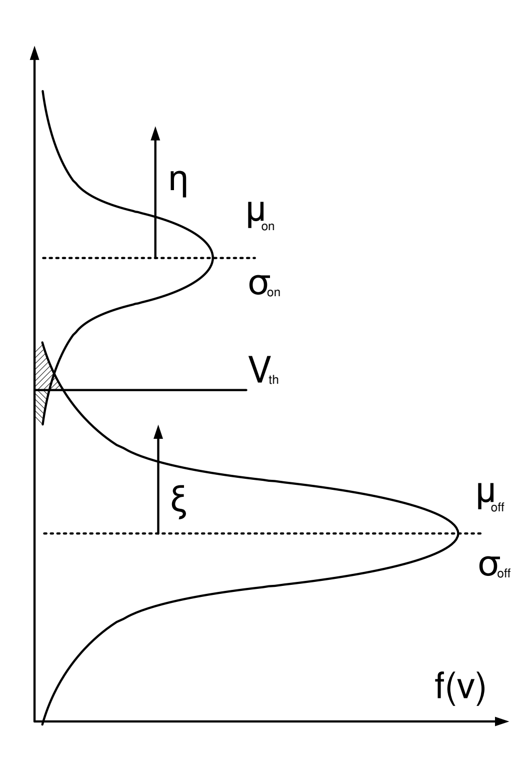


Figura 10.16: funciones de densidad de probabilidad de error y umbral de decisión.

probabilidad de decidir, también equivocadamente, que hemos recibido un "0" por medio de P_1. Calculemos la probabilidad de error total:

$$
\begin{aligned}
P_E &= aP_1 + (1-a)P_\circ \\
P_\circ &= \int_{V_{th}}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_{off}} e^{-\frac{(V-\mu_{off})^2}{2\sigma_{off}^2}} dV \\
P_1 &= \int_{-\infty}^{V_{th}} \frac{1}{\sqrt{2\pi}\sigma_{on}} e^{-\frac{(V-\mu_{on})^2}{2\sigma_{on}^2}} dV
\end{aligned}
\tag{10.28}
$$

Ahora si hacemos los siguientes cambios de variables: $\xi := \frac{V-\mu_{off}}{\sqrt{2}\sigma_{off}}$ y $\eta := \frac{V-\mu_{on}}{\sqrt{2}\sigma_{on}}$ y tenemos en cuenta que las gaussianas tienen simetría par con respecto a la media, resulta:

$$
\begin{aligned}
P_E &= aP_1 + (1-a)P_\circ \\
P_\circ &= \frac{1}{\sqrt{\pi}} \int_{Q_\circ}^{\infty} e^{-\xi^2} d\xi = \frac{1}{2} erfc(Q_\circ) \\
P_1 &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{-Q_1} e^{-\eta^2} d\eta = \frac{1}{\sqrt{\pi}} \int_{Q_1}^{\infty} e^{-\eta^2} d\eta = \frac{1}{2} erfc(Q_1)
\end{aligned}
\tag{10.29}
$$

Por tanto, deducimos:

$$P_E = \frac{a}{2} erfc[Q_\circ(V_{th})] + \frac{1-a}{2} erfc[Q_1(V_{th})] \tag{10.30}$$

Evidentemente, tomaremos como valor de V_{th} aquél que minimice la propabilidad de error:

$$\frac{dP_E}{dV_{th}} = \left(\frac{\partial P_E}{\partial Q_\circ}\frac{dQ_\circ}{dV_{th}}\right)_{V_{th}^*} + \left(\frac{\partial P_E}{\partial Q_1}\frac{dQ_1}{dV_{th}}\right)_{V_{th}^*} = 0 \Rightarrow -\frac{a}{\sqrt{2\pi}\sigma_{off}} e^{-Q_\circ^2} + \frac{1-a}{\sqrt{2\pi}\sigma_{on}} e^{-Q_1^2} = 0 \Rightarrow$$
$$\frac{a}{1-a}\frac{\sigma_{on}}{\sigma_{off}} = e^{Q_\circ^2 - Q_1^2} \Rightarrow \left(\frac{V_{th}^* - \mu_{off}}{\sqrt{2}\sigma_{off}}\right)^2 - \left(\frac{\mu_{on} - V_{th}^*}{\sqrt{2}\sigma_{on}}\right)^2 = \ln\left(\frac{a}{1-a}\frac{\sigma_{on}}{\sigma_{off}}\right) \tag{10.31}$$

Esta ecuación algebraica habría que resolverla numéricamente en cada caso particular para obtener el valor V_{th}^*. El caso descrito en casi cualquier libro de comunicaciones digitales es $a = \frac{1}{2}$ y $\sigma_{on} = \sigma_{off}$ que, sustituido en el que la expresión anterior lleva a [3]–[83]:

$$\begin{aligned} Q_\circ &= Q_1 \\ V_{th}^* &= \tfrac{\mu_{on}+\mu_{off}}{2} \\ P_E &= \tfrac{1}{2} erfc\left(\frac{\mu_{on} - \mu_{off}}{2\sqrt{2}\sigma}\right) \end{aligned} \tag{10.32}$$

10.9. Apéndice I: Deducción de los tiempos de respuesta interna.

Estimemos el tiempo que tarda un electrón en atravesar la zona p de entrada. Para hacerlo tenemos en cuenta que el flujo es puramente difusivo (por ser el portador minoritario). Por tanto, la expresión del flujo la podemos aproximar mediante:

$$\phi_n = D_n \frac{\partial \Delta n}{\partial x} \simeq D_n \frac{\Delta n}{w_p} \tag{10.33}$$

donde Δn representa la concentración de portadores en desequilibrio justo a la entrada del dispositivo [25] y w_p representa la anchura de la zona p.

A partir del flujo difusivo es fácil deducir la velocidad media de desplazamiento de los electrones. Para ello basta dividir el flujo por la concentración media de electrones ya que este cociente tienen las dimensiones físicas de una velocidad:

$$\frac{\phi_n}{\Delta n/2} \simeq \frac{w_p}{\tau_{dif}} \simeq \frac{2D_n}{w_p} \tag{10.34}$$

de donde deducimos que el tiempo medio de difusión de los electrones viene dado por la expresión siguiente:

$$\tau_{dif} = \frac{w_p^2}{2D_n} \tag{10.35}$$

[25]La aproximación hecha en (10.33) es muy buena puesto que la zona p es muy estrecha con respecto a la longitud de difusión y, en consecuencia, la curva exponencial se aproxima muy bien mediante una línea recta. La concentración de portadores en desequilibrio en la región de deplexión es pequeña, de ahí que la pendiente (negativa) de la recta se aproxime por el cociente $\frac{\Delta n}{w_p}$.

10.10. Problemas

Problema 12: Calcule la potencia óptica requerida en una comunicación en la que se transmiten datos con un régimen binario (R_b) de 155Mbps ($\Delta f = 310MHz$).
La longitud de onda de trabajo es $\lambda_\circ = 1300nm$. Se ha elegido un preamplificador cuyos parámetros de ruido son $(\Delta i_a^2)^{\frac{1}{2}} = 2pA/(Hz)^{\frac{1}{2}}$, $(\Delta v_a^2)^{\frac{1}{2}} = 2nV/(Hz)^{\frac{1}{2}}$. La capacidad total del circuito de entrada es C=3pF, la resistencia de carga es $R_L = 40K\Omega$ y la resistencia de realimentación (se trata de un amplificador de transimpedancia) es $R_f = 10K\Omega$. Se precisa en el receptor una K=20.
Datos: $q = 1{,}6 \times 10^{-19}$C, kT=0.026eV, $c_\circ = 3 \times 10^8 m/s$, $h = 6{,}6 \times 10^{-34} Js$.
1.- Considere un fotodiodo PIN con $\eta = 0{,}5$ y calcule la potencia óptica que de llegar al receptor, así como el número de fotones.
2.- Considere, ahora, un fotodiodo APD con $\eta = 0{,}2$. M=20 y $F = M^x = 10$ y determine la potencia óptica que ha de llegar al receptor así como el número de fotones.

Solución:

Apartado 1:

$$\left.\begin{array}{l}
\Delta i_a^2 = 4 \times 10^{-24} \frac{A^2}{Hz} \\[2ex]
\dfrac{\Delta v_a^2}{R^2} = \dfrac{6,25 \times 10^{-18}}{(8 \times 10^3)^2} = 6,25 \times 10^{-26} \dfrac{A^2}{Hz} \\[2ex]
\Delta v_a^2 \frac{4\pi^2}{3} C^2 (\Delta f)^2 = 4,54 \times 10^{-23} \frac{A^2}{Hz} \\[2ex]
\dfrac{4KT}{RM} = 2,08 \times 10^{-24} \dfrac{A^2}{Hz} \\[2ex]
2qF\bar{i} \simeq 0
\end{array}\right\} \Rightarrow \quad \frac{\bar{i}}{[5,16 \times 10^{-23}]^{1/2} [310 \times 10^6]^{1/2}} = 20 \Rightarrow$$

$$\bar{i} \ = \ 2,53 \mu A$$

con el valor de $\bar{i}$ obtenido se verifica que, efectivamente, $2qF\bar{i} = 8,09 \times 10^{-25} \frac{A^2}{Hz} \simeq 0$. Si además tenemos en cuenta que en una codificación NRZ el tiempo de definición de un bit es $T_b = \frac{1}{R_b}$, nos queda:

$$P_{opt} = \frac{\bar{i} \dfrac{hc_\circ}{\lambda}}{q\eta} = 4,81 \mu W$$

$$\overline{n} = \frac{P_{opt}}{hfR_b} = 203747$$

Apartado 2

$$\left.\begin{array}{r}
\dfrac{\Delta i_a^2}{M^2} = 10^{-26}\dfrac{A^2}{Hz} \\
\dfrac{\Delta v_a^2}{R^2M^2} = 1,56\times 10^{-26}\dfrac{A^2}{Hz} \\
\dfrac{\Delta v_a^2}{M^2}\dfrac{4\pi^2}{3}C^2(\Delta f)^2 = 1,135\times 10^{-25}\dfrac{A^2}{Hz} \\
\dfrac{4KT}{RM^2} = 5,2\times 10^{-27}\dfrac{A^2}{Hz} \\
2qF\overline{i} = 3,2\times 10^{-18}\overline{i}\dfrac{A^2}{Hz}
\end{array}\right\} \Rightarrow 20 = \frac{\overline{i}}{[3,2\times 10^{-18}\overline{i} + 1,28\times 10^{-25}]^{1/2}[310\times 10^6]^{1/2}}$$

La ecuación de segundo grado resultante es:

$$\overline{i^2} - 3,968\times 10^{-7}\overline{i} - 1,584\times 10^{-14} = 0 \Rightarrow \overline{i} = 4,33\times 10^{-7}A; \overline{Mi} = 8,66\mu A$$

con lo que resulta:

$$P_{opt} = \frac{\overline{i}\dfrac{hc_\circ}{\lambda}}{q\eta} = 824nW$$

$$\overline{n} = \frac{P_{opt}}{hfR_b} = 34905$$

Problema 13: Se tiene un sistema de comunicaciones analógicas en el que se quieren transmitir 30 canales de televisión utilizando 30 subportadores eléctricas, de manera que la señal resultante tiene una ocupación espectral de 800MHz. En el receptor que hay al otro extremo de la fibra, la capacidad parásita conjunta del fotodetector y el preamplificador es de 1pF. Para evitar posible fenómenos de saturación en un hipotético ecualizador se ha decidido emplear un preampificador de tensión, desechándose la opción del amplificador de transimpedancia y de un ecualizador en caso de que la primera etapa amplificadora filtre la señal. Teniendo esto en cuenta, se pide:
1.- Determine el valor de la resistencia de carga del fotodetector.
2.- Si los parámetros de ruido del amplificador son 1pA/$\sqrt{Hz}$ y 0,1nV/$\sqrt{Hz}$ compruebe que el término de ruido térmico dominante es el de la resistencia de carga.
3.- En las condiciones anteriores, calcule el valor óptimo de la ganacia interna del fotodetector.
4.- En la expresión del apartado anterior, tome x=1 (valor típico para un APD de germanio) y calcule la expresión de la SNR óptima.
5.- Teniendo en cuenta que en una comunicación analógica se necesita una SNR por encima de 40dB, razone si tiene sentido utilizar un APD de germanio como detector.
6.- Considere un fotodetector formado por el conjunto de un amplificador óptico más un fotodetector PiN y se modélelo como si fuese un APD con un valor muy bajo de x, x=0,01. Calcule la ganacia del amplificador óptico y la corriente necesaria en el fotodetector para tener una SNR de 40dB.
7.- Diseñe el amplificador óptico. $\alpha_\circ = \Gamma\sigma(N_2 - N_1)$ $\Gamma = 0,1$
$\sigma = 5 \times 10^{-25} m^{-3}$ $N_2 - N_1 = 2 \times 10^{24} m^{-3}$

Solución:

Apartado 1:

Como hay que emplear un preamplificador de tensión sin ecualización, el circuito RC de entrada no debe filtrar la señal:

$$\tilde{V} = \tilde{I} Z_{eq} = \frac{\tilde{I}}{Y_{eq}}$$

$$Y_{eq} = \frac{1}{R} + j\omega C_p = \frac{1 + j\omega R C_p}{R} \Rightarrow Z_{eq} = \frac{R}{1 + j\omega R C_p}$$

$$\omega R C_p = 1 \Rightarrow R_{max} = \frac{1}{2\pi f_{max} C_p} = \frac{1}{2\pi \times 800 \times 10^6 \times 10^{-12}} \simeq 200\Omega$$

Apartado 2:

Calculemos cada uno de los términos de ruido cuántico:

$$\frac{4KT}{R} = \frac{4 \times 0,026 \times 1,6 \times 10^{-19}}{200} = 8 \times 10^{-23} \frac{A^2}{Hz}$$

$$\Delta i_a^2 = 10^{-24} \frac{A^2}{Hz}$$

$$\frac{\Delta v_a^2}{R^2} = 2,5 \times 10^{-25} \frac{A^2}{Hz}$$

$$\frac{4\pi^2 \Delta v_a^2 C^2 (\Delta f)^2}{3} = 8,41 \times 10^{-23} \frac{A^2}{Hz}$$

Luego, efectivamente, el ruido térmico dominante es el de la resistencia de carga.

Apartado 3:

Tenemos que maximizar la expresión de la SNR:

$$\left.\begin{array}{l} K^2 \simeq \dfrac{\overline{i^2}}{\left(2q\overline{M}^x \overline{i} + \dfrac{4kT}{R\overline{M}^2}\right)\Delta f} \\ \dfrac{dK^2}{d\overline{M}} = 0 \end{array}\right\} \Rightarrow 2q\overline{i}x\overline{M}^{x-1} - \frac{2\overline{M}}{\overline{M}^4}\frac{4KT}{R} = 0 \Rightarrow 2qx\overline{M}^{x-1} = \frac{4KT}{R}\frac{2}{\overline{M}^3} \Rightarrow$$

$$\overline{M}_{opt} = \left(\frac{4KT}{xRq\overline{i}}\right)^{\frac{1}{2+x}} \Rightarrow \overline{M}_{opt} = \left(\frac{5 \times 10^{-4}}{x\overline{i}}\right)^{\frac{1}{2+x}}$$

Apartado 4:

En la expresión anterior tomamos x=1 y sustituimos en la fórmula de la SNR:

$$\overline{M}_{opt} = \sqrt[3]{\frac{4KT}{qR\overline{i}}}$$

$$SNR]_{opt} = K_{opt}^2 = \frac{\overline{i^2}}{\left[2q\left(\frac{4KT}{Rq\overline{i}}\right)^{\frac{1}{3}}\overline{i} + \frac{4kT}{R(\frac{4KT}{Rq\overline{i}})^{\frac{2}{3}}}\right]\Delta f} = \frac{\overline{i}^{\frac{4}{3}}}{\left[\left(\frac{32q^2kT}{R}\right)^{\frac{1}{3}} + \left(\frac{4kTq^2}{R}\right)^{\frac{1}{3}}\right]\Delta f} =$$

$$= \frac{\overline{i}^{\frac{4}{3}}}{3\Delta f\left(\frac{4q^2kT}{R}\right)^{\frac{1}{3}}}$$

Si sustituimos valores queda:

SNR	$\bar{i}^{\frac{4}{3}}$	$\overline{M}_{opt}$
40dB=10^4	12,97μA	6,2
50dB=10^5	72,94μA	2,6
40dB=10^6	410,19μA	1,1

$$(SNR)_{opt} = \frac{\bar{i}^{\frac{4}{3}}}{3,0478 \times 10^{-11}}$$

Apartado 5:

Imponiendo una SNR de 40, 50 y 60dB y teniendo en cuenta que $\bar{i} = (3,0478 \times 10^{-11})^{\frac{3}{4}}$ y $M_{opt} = \sqrt{\frac{5\times10^{-4}}{\bar{i}}}$ obtenemos los datos de la tabla (10.10). Vemos que las ganacias óptimas son muy bajas, debido al elevado factor de ruido del fotodetector. Las gananacia interna de un APD de germanio es del orden de 100, por lo que no tiene sentido emplearlo en esta aplicación.

Apartado 6:

Como ahora tenemos x=0,01, la sensibilidad del receptor va a ser más alta. Calculémosla:

$$M_{opt} = \left(\frac{4KT}{xRq\bar{i}}\right)^{\frac{1}{2+x}} \simeq \left(\frac{4KT}{xRq\bar{i}}\right)^{\frac{1}{2}}$$

$$SNR_{opt} = \frac{\bar{i}^2 M_{opt}^2}{\left(2q\bar{i}\overline{M}^{2+x} + \frac{4KT}{R}\right)\Delta f} \simeq \frac{\bar{i}^2 \frac{4KT}{xRq\bar{i}}}{\left(2q\bar{i}\frac{4KT}{xRq\bar{i}} + \frac{4KT}{R}\right)\Delta f} = \frac{\bar{i}}{q\Delta f(x+2)} \Rightarrow$$

$$\bar{i} = SNR_{opt} qB(x+2) = 10^4 \times 1,6 \times 10^{-19} \times 8 \times 10^8 \times 2,01 = 2,57\mu A$$

$$M_{opt} \simeq \sqrt{\frac{4KT}{xqR\bar{i}}} = 139,5$$

Apartado 7:

Diseñemos el amplificador óptico. Para ello, tendremos en cuenta que $G_{EDFA} = e^{\alpha_\circ L}$ y que $\alpha_\circ = \Gamma\sigma(N_2 - N_1)$

$$L = \frac{\ln G}{\Gamma\sigma(N_2 - N_1)} = \frac{\ln 139,5}{0,1 \times 5 \times 10^{-25} \times 2 \times 10^{24}} = 49,37m$$

Capítulo 11

Amplificadores ópticos

Resumen: En este capítulo vamos a explicar el funcionamiento del amplificador óptico de fibra dopada con erbio (EDFA). En los apartados dos, tres y cuatro haremos una aproximación que, siendo admisible en los amplificadores ópticos a semiconductor (SOAs), no lo es en los EDFAs. Esta aproximación consiste en suponer que la inversión de población del medio activo es la misma en todos los puntos del amplificador. De esta manera resulta sencillo introducir los conceptos de ancho de banda de amplificación, saturación de la ganancia y factor de ruido del amplificador. En los dos últimos apartados modelaremos el comportamiento del EDFA en sus dos configuraciones básicas: la de copropagación y la de contrapropagación. Un aspecto no tenido en cuenta a la hora de deducir este modelo es que las secciones cruzadas asociadas a la emisión estimulada y a la absorción son distintas y, a su vez, estas dos secciones cruzadas dependen de la frecuencia a lo largo de todo el ancho de banda de operación. Al no haber tenido en cuenta estos aspectos, lo modelos obtenidos tan sólo reproducen cualitativamente el funcionamiento de los amplificadores de fibra dopada con erbio.

11.1. Introducción

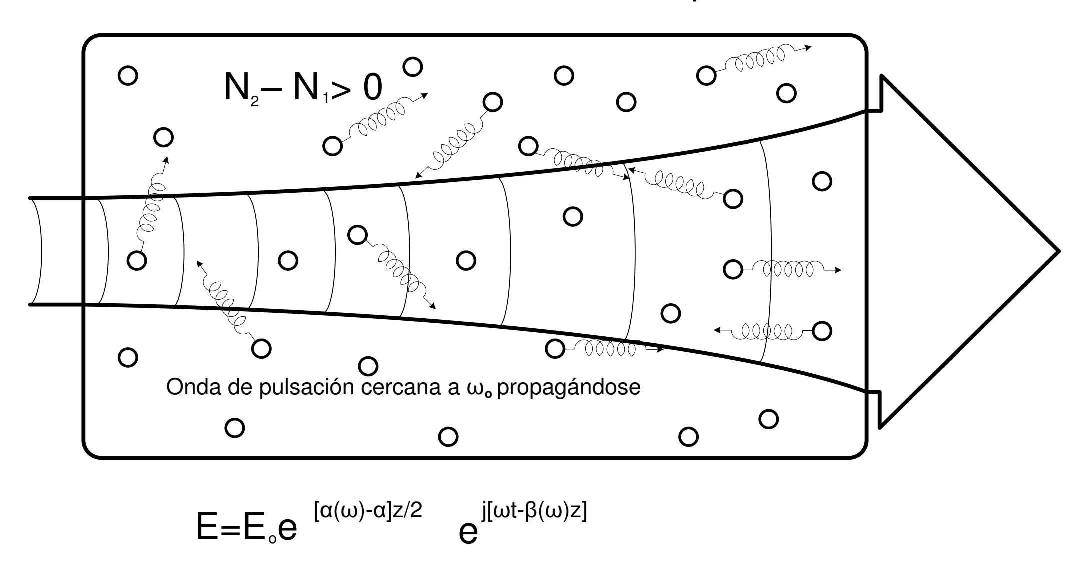


Figura 11.1: Medio atómico con inversión de población que ampifica la radiación electromgnética incidente.

En las comunicaciones por fibra óptica se utilizan dos tipos de amplificadores ópticos: 1.- El EDFA [13]–[16]–[27] (del inglés "Erbium Doped Fiber Amplifier") cuyo medio activo son iones de erbio con carga positiva trivalente que, inmersos en la matriz vítrea del núcleo de una bobina de fibra monomodo, tienen estados energéticos accesibles discretos. 2.- El SOA [25] (del inglés "Semiconductor Optical Amplifier") es un amplificador óptico cuyo medio activo es un semiconductor. El EDFA es, con diferencia, el más utilizado de los dos, ya que se emplean sistemáticamente como amplificador de línea [72], como "booster" [72] y como preamplificador [72] para aumentar la sensibilidad de los receptores. El SOA es un dispositivo con muchas posibles aplicaciones en redes todo–ópticas que, quizás, se desarrollen en un futuro no muy lejano. Sus aplicaciones más prometedoras son los convertidores en longitud de onda [25]–[59]–[58], la realización física de regeneradores 3R [25]–[59]–[58] y de dispositivos capaces de llevar a cabo una multiplexación temporal de pulsos en el dominio óptico [25]–[59]–[58]. Su funcionamiento se basa en una serie de efectos no lineales que, normalmente, están inducidos por la variación de la concentración de los portadores libres en el medio activo, (aunque, dicho sea de paso, la dinámica asociada a estos fenómenos es relativamente lenta, del orden de 50ps, pero se cree que la nueva generación de SOAs basados en puntos cuánticos permitirá superar este inconveniente). En este capítulo nos centraremos en el estudio del EDFA ya que éstos son componentes comerciales que se emplea sistemáticamente en las redes ópticas actuales.

En capítulos anteriores estudiamos los mecanismos de interacción luz–materia en

medios cuyos estados energéticos accesibles estaban discretizados, como el mostrado en la fig. {11.1}. Recordemos una serie de resultados que obtuvimos entonces:

$$\left.\begin{aligned} \frac{d\phi}{dz} &= \alpha_\circ(f)\phi + \xi_{es}(f) \\ \alpha_\circ(f) &= \Gamma\sigma(f)(N_2 - N_1) \\ \sigma(f) &= \frac{\lambda^2 l(f)}{8\pi\tau_{sp}} \\ l(f) &= \frac{2}{\pi\Delta f}\frac{1}{1+\left(\dfrac{f-f_{21}}{\Delta f}\right)^2} \end{aligned}\right\} \tag{11.1}$$

La primera ecuación de (11.1) describe la evolución del flujo ϕ de una onda electromagnética que se propaga por un medio activo. El término $\alpha_\circ(f)$ recibe el nombre de ganancia por unidad de longitud puesto que los tres términos que intervienen en la ecuación tienen, necesariamente, las mismas unidades físicas. Dado que $\frac{d\phi}{dz}$ es, dimesionalmente, un flujo dividido por una longitud, $\alpha_\circ$ ha de tener la dimensión de la inversa de una longitud y de ahí su nombre.

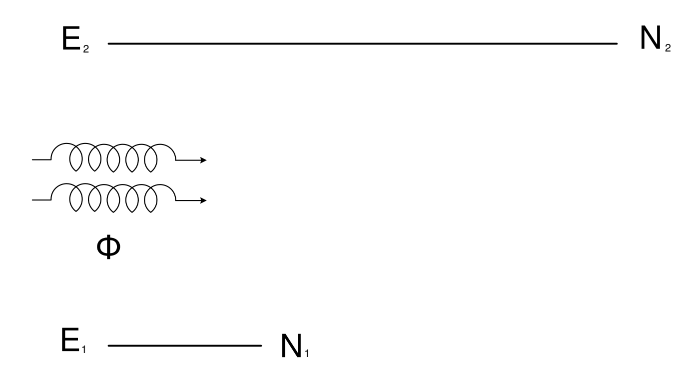


Figura 11.2: Estados energéticos accesibles del medio dieléctrico.

La segunda fórmula de (11.1) indica que la ganancia por unidad de longitud del medio amplificador, $\alpha_\circ(f)$, es positiva sólo si el número de átomos (por unidad de volumen) [1] que hay en el estado energético excitado, N_2, es mayor que el que hay en el estado inferior, N_1. El término Γ, que interviene en la definición de la ganacia por unidad de longitud, recibe el nombre de factor de confinamiento y pone de manifiesto que sólo una fracción del flujo (esta fracción, expresada en tanto por uno, es, precisamente, el factor de confinamiento) se propaga por el medio activo e interacciona con éste. El término $\sigma(f)$ recibe el nombre de sección cruzada debido a que tiene unidades de superficie. De alguna manera, si interpretamos los procesos de interacción luz–materia como la colisión de un fotón con un ión de erbio, la sección cruzada es el área que presenta el ión de erbio (considerado como una "diana")

[1] Suponemos el mismo grado de degeneración en los dos estados energéticos accesibles.

frente a un flujo incidente de fotones (que, en esta inexacta analogía, serían los "proyectiles"). Cuanto mayor es la sección cruzada, más intensos son los procesos de interacción luz–materia. Por último, $l(f)$ recibe el nombre de función de línea y su significado es claro: representa la densidad de probabilidad de la variable aleatoria asociada a que un fenómeno de interacción luz–materia se produzca en un intervalo de frecuencias *df* determinado. En (11.1) hemos supuesto que la función de línea es una Lorentziana lo cual es cierto sólo en medios activos de niveles discretos y ensanchamiento homogéneo (que no es el caso de los iones de erbio del EDFA). Por último, el término $\xi_{es}(f)$ representa el pequeño incremento de flujo asociado a la fracción de emisión espontánea que se propaga en la misma dirección y sentido y en el mismo ancho de banda que el flujo incidente.

En la fig. {11.2} observamos el diagrama de estados energéticos accesibles de un medio activo en el que hay inversión de población. El flujo incidente, que está compuesto por fotones de frecuencia $f_{21} = \frac{E_2 - E_1}{h}$, estimula procesos de *absorción* y *emisión estimulada.* Como $N_2 > N_1$, la emisión estimulada es más importante que la absorción y el flujo crece exponencialmente a medida que se propaga por el medio, tal como se muestra en la fig. {11.1}.

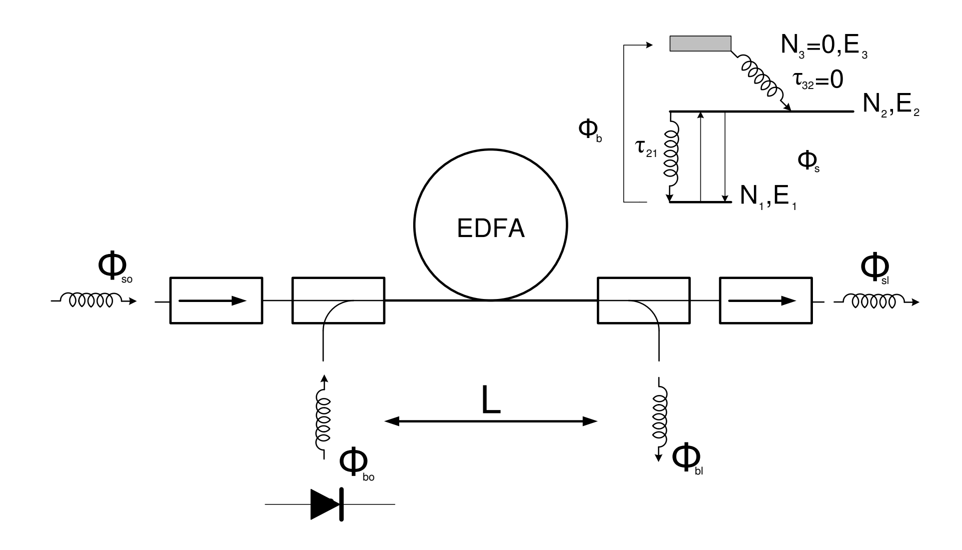


Figura 11.3: Amplificador óptico de fibra dopada con erbio. Bombeo en copropagación.

Los amplificadores ópticos de fibra dopada con erbio [2] tienen una estructura parecida a la mostrada en la fig. {11.3}. Consisten en una bobina de fibra óptica monomodo cuya longitud es típicamente de unas cuantas decenas de metros, en cuyo núcleo se introduce una cierta concentración de iones de erbio, una tierra rara, que denotamos Er^{3+}. Estos iones constituyen el medio activo del amplificador óptico y sus estados energéticos accesibles se muestran también en la fig. {11.3}. Este diagrama es el típico de un láser de tres niveles discretos, en el que el estado energético "3" constituye la banda de bombeo y los estados

[2]Se fabrican también amplificadores de fibra dopada con yterbio para enlaces de segunda ventana. Esto dispositivos no han llegado, sin embargo, a popularizarse por una serie de inconvenientes técnicos que no vamos a explicar.

"1" y "2" son los estados entre los que se producen las transiciones estimuladas (absorción y emisión estimulada). Como en cualquier láser de tres niveles, el estado "3" es inestable (es decir, el tiempo medio de permanencia de un ión en este estado es muy pequeño, $\tau_{32} \simeq 0$), el nivel "2" es metaestable (es decir, τ_{21} es grande) y, por último, el estado "1" es estable ya que es el estado de mínima energía).

Por el amplificador óptico se propagan dos flujos de radiación luminosa de longitudes de onda distintas. Por un lado está el flujo de bombeo (que normalmente está generado por un láser semiconductor) y que, en en la fig. {11.3}, se propaga en el mismo sentido que la otra radiación luminosa, la de la señal que se quiere amplificar. A esta configuración de bombeo, por motivos obvios, se le da el nombre de bombeo en copropagación. También es posible realizar el bombeo desde el otro extremos de la fibra y en sentido contrario, en cuyo caso tendríamos una configuración en contrapropagación. Se pueden utilizar dos longitudes de onda distintas para realizar el bombeo: 980nm y 1480nm, ya que el Er^{3+} tiene estas dos bandas de absorción. Las características del EDFA (mayor o menor sensibilida a la saturación y factor de ruido más o menos acusado) dependen tanto del tipo de configuración de bombeo (copropagación, contrapropagación o bombeo bidireccional) como de la banda en la que se efectúe el bombeo; desgraciadamente, el modelado que vamos a hacer del EDFA no permitirá predecir estas caracteríticas. El motivo es que tendríamos que estudiar previamente la espectroscopía de las tierras raras [13] (el erbio lo es) y el ensanchamiento que produce la matriz vítrea sobre los estados energéticos, lo que claramente excede los objetivos marcados en este libro. Destaquemos, por último, que los EDFAs tienen una ancho de banda de unos 30nm en los que la ganancia es prácticamente plana [3].

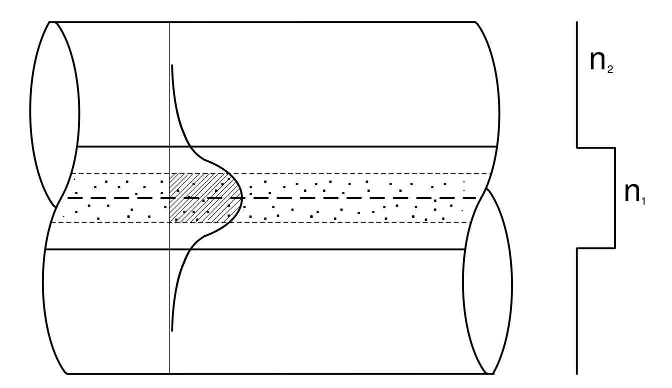


Figura 11.4: Factor de confinamiento de la fibra dopada con erbio.

En la fig. {11.4} se ha representado un trozo de la fibra óptica monomodo dopada con erbio. Vemos que el dopado se realiza en el núcleo de la fibra pero no necesariamente en toda su sección. El significado del factor de confinamiento Γ es claro: es el área rallada dividida por el área total, correspondiente a la fracción de intensidad luminosa que se propaga por

[3]Esto se consigue introduciendo en el EDFA una fibra con pérdidas que ecualiza la ganancia en el rango de frecuencias de interés.

la zona activa de la fibra expresada en tanto por uno. Sólo esta fracción produce fenómenos de interacción luz–materia. Puesto que el flujo de bombeo y el de la señal se realizan a longitudes de onda distintas, los factores de confinamiento asociados son, también, diferente. Con objeto de no complicar el modelado del EDFA los tomaremos iguales aunque esto no es, evidentemente, cierto.

Hay otro aspecto de los EDFAs que no vamos a estudiar: los iones trivalentes del erbio presentes en el núcleo de la fibra generan un campo eléctrico que influye sobre sus estados energéticos accesibles. Este campo eléctrico hace que, por efecto Stark [14], estos estados accesibles se separen en bandas que son aún más anchas, modificando las secciones cruzadas del medio activo. A efectos del modelado del EDFA, la consecuencia más importante es que la curva que relaciona la sección cruzada de la emisión estimulada frente a la frecuencia de la radiación luminosa es distinta de la de la absorción [13]–[16]–[27]–[36]. Como ya hemos explicado anteriormente, el estudio de la espectroscopía de las tierras raras, el grado de degeneración de los estados energéticos accesibles y la influencia que tiene en éstos la matriz vítrea (el sílice) y el efecto Stark rebasa, con creces, los límites de este libro por lo que vamos a ignorarlos y tomaremos secciones cruzadas iguales para los fenómenos de emisión estimulada y absorción. Tampoco vamos a intentar modelar de una manera precisa el ruido de emisión espontánea (ruido "ASE") que se produce en los EDFAs. Pedimos disculpas por esta grosera simplificación de la realidad y remitimos al lector interesado en profundizar en estas interesantes cuestiones a la bibliografía adjunta [13]–[14]–[16]–[27]–[36].

En los tres apartados que siguen haremos una aproximación que, siendo relativamente válida en los amplificadores ópticos a semiconductor (salvo la expresión de la función de línea) [25]–[66]–[86], no se puede justificarse en los amplificadores de fibra dopada con erbio. Esta aproximación consiste en suponer que la inversión de población del medio activo es la misma en todos los puntos del amplificador. Lo que pretendemos con esto es introducir de una manera sencilla los conceptos de anchos de banda de amplificación (ancho de banda asociado a la ganancia por unidad de longitud y el asociado a la ganancia total), el fenómeno de saturación de la ganancia y el factor de ruido de un amplificador óptico. Creemos que, puesto que estos fenómenos se dan en todos los amplificadores ópticos, el "insight" que se adquiere con este tratamiento tan sencillo compensa la falta de rigor.

11.2. Ganancias y anchos de banda asociados

El primer ancho de banda que podemos definir en un amplificador óptico está relacionado con la ganancia por unidad de longitud. Lo definimos como el intervalo de frecuencias para el que esta ganancia es superior o igual a la mitad del máximo y los denotamos mediate $FWHM_\alpha$ (del inglés "full width half maximum"). En la fig. {11.5} hemos representado la dependencia de la ganancia por unidad de longitud asociada a una función de línea Lorentziana (dividida por su valor máximo) frente a la frecuencia y hemos trazado también una recta horizontal de valor $\frac{1}{2}$. Los dos puntos de corte entre definen $FWHM_\alpha$. Veamos cuál es su expresión analítica:

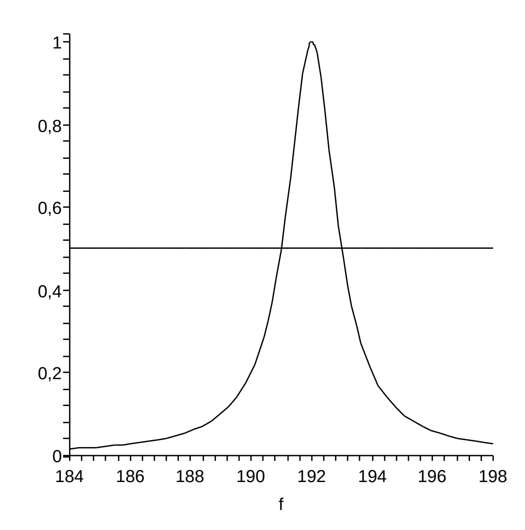


Figura 11.5: Ancho de banda asociado a la ganancia por unidad de longitud. En el eje vertical se ha representado $\frac{\alpha_\circ(f)}{\alpha_{\circ max}}$ junto a una línea horizontal de ordenada 0,5 y en el horizontal la frecuencia medida en THz. El valor de f_{21} es $f_{21} = 192THz$.

$$\left.\begin{array}{l} \alpha_\circ(f) = \dfrac{\Gamma\lambda^2(N_2 - N_1)}{8\pi\tau_{es}} \dfrac{\dfrac{2}{\pi\Delta f}}{1 + (\dfrac{f - f_{21}}{\Delta f})^2} \\ \\ \alpha_{\circ max} = \alpha_\circ(f_{21}) = \dfrac{2}{\pi\Delta f}\dfrac{\Gamma\lambda^2(N_2 - N_1)}{8\pi\tau_{es}} \end{array}\right\} \Rightarrow \frac{\alpha_{\circ max}}{2} = \alpha_\circ \Rightarrow f = f_{21} \pm \frac{\Delta f}{2} \tag{11.2}$$

Por tanto el ancho de banda resulta ser:

$$FWHM_\alpha = \Delta f \tag{11.3}$$

Determinemos, a continuación, la ganancia total del amplificador óptico teniendo en cuenta que ésta es, por definición, el cociente entre el flujo de salida y el flujo de entrada:

$$\left.\begin{array}{c} \dfrac{d\Phi}{dz} = \alpha_\circ(f) \\ \\ z = 0 \ \Phi = \Phi_\circ; \ z = L \ \Phi = \Phi_L \end{array}\right\} \Rightarrow \int_{\Phi_\circ}^{\Phi_L} \frac{d\Phi}{\Phi} = \int_0^L \alpha_\circ(f)dz \Rightarrow G(f) = e^{\alpha_\circ(f)L} \tag{11.4}$$

La ganancia total máxima se produce también para la frecuencia $f = f_{21}$:

$$G_{\circ max} = G_\circ(f = f_{21}) = e^{\alpha_{\circ max}L} \tag{11.5}$$

Si ahora imponemos la condición de que sea igual a la mitad del máximo, resulta:

$$\left.\begin{array}{r}\dfrac{e^{\alpha_\circ(f)L}}{G_{\circ max}}=\dfrac{1}{2};\ \ln\dfrac{G_{\circ max}}{exp\left(\dfrac{\alpha_{\circ max}L}{1+(\frac{f-f_{21}}{\Delta f})^2}\right)}=\ln 2;\\ \ln G_{\circ max}-\dfrac{\ln G_{\circ max}}{1+\left(\dfrac{f-f_{21}}{\Delta f}\right)^2}=\ln 2;\\ \left(\dfrac{f-f_{21}}{\Delta f}\right)^2=\dfrac{\ln G_{\circ max}}{\ln G_{\circ max}/2}-1\end{array}\right\}\Rightarrow FWHM_G=\Delta f\sqrt{\frac{\ln 2}{\ln G_{\circ max}/2}} \quad (11.6)$$

Puesto que normalmente $G_\circ \gg 1$ resulta que $FWHM_G \ll FWHM_\alpha$.

11.3. Saturación de la ganancia

Naturalmente un modelo que predice un crecimiento exponencial del flujo de salida tiene un rango de validez muy limitado (sólo predice resultados correctos si el flujo de entrada es muy débil). La ganancia debe disminuir por debajo de lo predicho por el modelo del apartado anterior a medida que el flujo se haga del mismo orden de magnitud o mayor que un determinado valor umbral que llamaremos *flujo de saturación. Postulamos*, como modelo más relista del EDFA, uno en el que la ganancia por unidad de longitud decrece con respecto al flujo de acuerdo con la siguiente expresión:

$$\alpha_\circ^s=\frac{\alpha_\circ(f)}{1+\dfrac{\Phi}{\Phi_{sat}}} \quad (11.7)$$

La ganancia por unidad de longitud es proporcional a la inversión de población N_2-N_1, pero si el flujo que atraviesa el amplificador se hace muy grande, el número de iones que pasan del estado excitado al fundamental aumenta, disminuyendo, por tanto, la inversión de población. La relación más sencilla que podemos establecer y que refleja esta disminución es (11.7). El parámetro Φ_{sat} es una característica propia del amplificador que hay que determinar experimentalmente con algún tipo de ensayo. La ecuación diferencial resultante, que describe la variación del flujo en un amplificador óptico con el modelo de saturación propuesto, es integrable:

$$\frac{d\Phi}{dz}=\frac{\alpha_\circ(f)}{1+\dfrac{\Phi}{\Phi_{sat}}}\Phi\Rightarrow\frac{d\Phi}{\Phi}+\frac{d\Phi}{\Phi_{sat}}=\alpha_\circ(f)dz\Rightarrow\int_{\Phi_{in}}^{\Phi_{out}}\frac{d\Phi}{\Phi}+\frac{\Phi_{out}-\Phi_{in}}{\Phi_{sat}}=\ln G_\circ(f) \quad (11.8)$$

La ganacia *total* del amplificador la vamos a denotar mediante $G(f)$, mientras que la ganancia sin saturación como $G_\circ(f)$. Teniendo esto en cuenta resulta:

$$\ln G(f) + \frac{\Phi_{out}}{\Phi_{sat}}[1 - \frac{1}{G(f)}] = \ln G_{\circ}(f) \Rightarrow G(f) = G_{\circ}(f) exp\left(-\frac{G(f)-1}{G(f)}\frac{\Phi_{out}}{\Phi_{sat}}\right) \quad (11.9)$$

La relación (11.9) define una ecuación implícita que nos permite calcular la ganancia del amplificador óptico para un valor determinado del flujo de de salida. Con objeto de tener una idea cualitativa del comportamiento del amplificador, fijamos la frecuencia al valor $f = f_{12}$:

$$G_{max} = G_{\circ max} exp\left(-\frac{G_{max}-1}{G_{max}}\frac{\Phi_{out}}{\Phi_{sat}}\right) \quad (11.10)$$

y, dado que para valores razonable de Φ_{in} (es decir, flujos de entrada muy inferiores al de saturación), se verifica $G_{max} \gg 1$ por lo que $\frac{G_{max}-1}{G_{max}} \simeq 1$ resultando:

$$G_{max} \simeq G_{\circ max}(f) exp\left(-\frac{\Phi_{out}}{\Phi_{sat}}\right) \quad (11.11)$$

Vemos que cuando el flujo de salida se hace del mismo orden de magnitud que el flujo de saturación, la ganancia del amplificador óptico decrece exponencialmente. ¿Qué tipo de ensayo habría que hacer para determinar el flujo de saturación Φ_{sat} de un amplificador óptico? Supongamos que introducimos inicialmente un flujo a la entrada que es lo suficientemente débil como para que la ganancia (es decir, el cociente entre el flujo de salida y el de entrada) sea igual a $G_{\circ}$. Supongamos, también, que aumentamos este flujo de manera que la ganancia del amplificador empieza a caer por debajo del valor teórico máximo $G_{\circ}$. Si continuamos aumentando la entrada hasta que $G \equiv \frac{\Phi_{out}}{\Phi_{in}} = \frac{G_{\circ}}{2}$, en este momento anotamos el flujo de salida y lo denotamos, por ejemplo, como Φ_{out}^{sat}. En estas condiciones tendremos:

$$\frac{G_{\circ max}}{2} = G_{\circ max} exp\left(-\frac{\frac{G_{\circ max}}{2}-1}{\frac{G_{\circ max}}{2}}\frac{\Phi_{out}^{sat}}{\Phi_{sat}}\right) \Rightarrow \Phi_{sat} = \frac{G_{\circ max}-2}{G_{\circ max}\ln 2}\Phi_{out}^{sat} \quad (11.12)$$

Esta relación nos permite, pues, calcular el valor de Φ_{sat} a partir de la medida experimental de Φ_{out}^{sat}

11.4. Factor de ruido de un amplificador óptico

En los apartados anteriores hemos integrado la ecuación diferencial asociada al flujo de la señal sin tener en cuenta el término de emisión espontánea. Sin embargo, una pequeña fracción de estos fotones se refleja en la superficie de separación entre núcleo y corteza, produciendo una señal a la salida en ausencia de entrada. Esta señal es, evidentemente, un

ruido que se superpone a cualquier radiación luminosa amplificada por el EDFA. Vamos a determinar la densidad espectral de potencia asociada a este ruido (por supuesto, en el marco del modelo simplificado que estamos considerando).

$$\left.\begin{array}{c}\dfrac{d\Phi}{dz} = \alpha_\circ(f)\Phi + \xi_{sp}(f) \\ \\ z = 0 \quad \Phi = 0\end{array}\right\} \tag{11.13}$$

Como queremos calcular la densidad espectral de potencia del ruido a la salida del amplificador, no consideramos ningún flujo a la entrada. Cualquier señal que aparezca a la salida estará provocada por los fotones generados por emisión espontánea y posteriormente amplificados en el dispositivo. La ecuación diferencial (11.13) es de variables separadas, de manera que la podemos integrar fácilmente:

$$\frac{d\Phi}{\alpha_\circ(f)\Phi + \xi_{sp}(f)} = dz \Rightarrow \ln\left[\alpha_\circ(f)\Phi + \xi_{sp}(f)\right]_0^{\Phi_L} = \alpha_\circ(f)L \Rightarrow$$

$$\Phi_L = \frac{\xi_{sp}(f)}{\alpha_\circ(f)}[e^{\alpha_\circ L} - 1]$$

Recordemos que ξ_{sp} es la fracción de fotones que, generados por emisión espontánea en el núcleo de la fibra, son guiados por la fibra en el mismo sentido que el flujo de la señal, caen en el mismo ancho de banda que ésta y tienen polarización vertical (suponemos que a la salida del amplificador hay un polarizador orientado según esta dirección de manera que podemos eliminar la mitad del ruido).

$$\xi_{sp}(f) = \frac{N_2}{\tau_{sp}} l(f) B_{el} \frac{d\Omega}{2 \times 4\pi} \tag{11.14}$$

- $\dfrac{N_2}{\tau_{sp}}$ es el número total de fotones generados por emisión espontánea por unidad de volumen y tiempo.
- $l(f)B_{el}$ es la fracción de fotones que coincide en ancho de banda con la señal (el resto lo podemos eliminar con un filtro óptico paso–banda).
- $\dfrac{d\Omega}{2 \times 4\pi}$ es la fracción de fotones espontáneos guidos por la fibra, que se propagan en el mismo sentido que la señal y tienen polarización vertical. $d\Omega$ es el ángulo sólido asociado a las direcciones para las que los fotones generados por emisión espontánea son guiados por la fibra.

La expresión del coeficiente de ganacia por unidad de longitud es:

$$\alpha_\circ(f) = \Gamma \frac{\lambda^2 l(f)}{4\pi\tau_{sp}}(N_2 - N_1) \tag{11.15}$$

Con lo que el cociente $\frac{\xi(f)}{\alpha_\circ(f)}$ tiene la siguiente expresión:

$$\frac{\xi(f)}{\alpha_\circ(f)} = \frac{\frac{N_2}{\tau_{sp}} l(f) B_{el} \frac{d\Omega}{2 \times 4\pi}}{\Gamma \frac{\lambda^2 l(f)}{4\pi\tau_{sp}} (N_2 - N_1)} = \frac{N_2}{N_2 - N_1} \frac{d\Omega}{\Gamma\lambda^2} B_{el} \tag{11.16}$$

La potencia a la salida del amplificador óptico es $P_L = \Phi_L A_t h f$ donde A_t es la sección efectiva de la fibra. Teniendo esto en cuenta resulta:

$$P_L = \frac{N_2}{N_2 - N_1} \frac{d\Omega}{\frac{\Gamma\lambda^2}{A_t}} (G_\circ - 1) h f B_{el} \tag{11.17}$$

Ahora tendríamos que demostrar que el ángulo sólido $d\Omega$ es precisamente $\frac{\Gamma\lambda^2}{A_t}$ (no lo vamos a hacer para no complicar más la demostración), con lo que queda:

$$P_L = \frac{N_2}{N_2 - N_1} \frac{d\Omega}{\frac{\Gamma\lambda^2}{A_t}} (G_\circ - 1) h f B_{el} \Rightarrow S_{ASE} = n_{sp}(G_\circ - 1) h f \simeq n_{sp} G_\circ h f \tag{11.18}$$

En la relación (11.18) comprobamos que si el amplificador funciona en conduciones ideales (es decir, con inversión de población completa) $n_{sp} = 1$, el ruido ASE (del inglés "Amplified Spontaneous Emission") sería, en este caso, equivalente a suponer que a la entrada, aparte de la señal a amplificar, hay, además, un fotón incorrelado con ésta.

Para definir el factor de ruido de un amplificador óptico hay que pasar, sin embargo, al dominio eléctrico ya que la detección introduce términos de "mezcla" entre la señal útil amplificada y el ruido ASE (esta interferencia entre la señal útil y el ruido ASE es, por supuesto, un ruido indeseado). El campo eléctrico total en el fotodetector es la suma del campo debido a la señal útil amplificada y el asociado al ruido ASE superpuesto. Dado que el fotodetector genera una corriente que es proporcional al cuadrado del campo eléctrico total, aparecen términos de interferencia entre las distintas señales electromagnéticas. Los términos de ruido dominantes son a) El ruido cuántico asociado a la señal amplificada y b) El ruido de interferencia (o mezcla) entre la señal amplificada y el ruido ASE. Esta afirmación es cierta siempre que se tenga la precaución de colocar a la salida de amplificador óptico un filtro pasobanda óptico que limite el ancho de banda del ruido ASE que llega al fotodetector. Si esto no se hace, dos términos de ruido que no estamos calculando, la mezcla del ruido ASE consigo mismo y el ruido cuántico asociado al ruido ASE, se hacen importantes e incluso pueden llegar a ser dominantes.

A la hora de calcular la figura de ruido del amplificador óptico supondremos, para simplificar el razonamiento, que la señal a amplificar es perfectamente monocromática tal y como la hemos representado en la fig. {11.6} (esta simplificación, aunque facilita los cálculos,

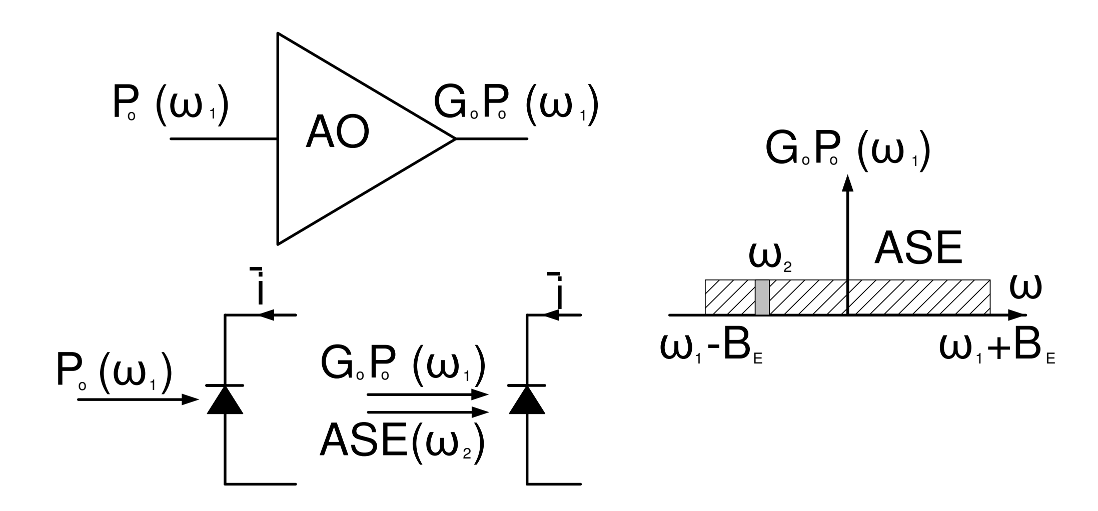


Figura 11.6: Cálculo de la figura de ruido de un amplificador óptico.

no es imprescindible). La figura de ruido (o factor de ruido) de un amplificador es, por definición, el cociente entre la relación señal a ruido a la entrada dividido por la relación señal a ruido a la salida del amplificador, es decir, $F := \frac{\frac{S}{N})_\circ}{\frac{S}{N})_L}$. Evidentemente, el cociente que resulta difícil de calcular es la relación señal a ruido a la salida del amplificador.

Cuando se produce la mezcla de la señal útil $G_\circ P_\circ$ con el ruido ASE (ver fig. {11.6}), las componentes espectrales situadas por encima de la pulsación $\omega_1 + B_E$ y por debajo de $\omega_1 - B_E$ caen, al mezclarse, fuera del ancho de banda de la primera etapa amplificadora del receptor, motivo por el que no vamos a tenerlas en cuenta. Consideremos una componente de ruido ASE arbitraria de pulsación $\omega_2 \in (\omega_2, \omega_2 + d\omega_2)$. La corriente generada al mezclarse en el fotodetector este ruido ASE con la señal útil $G_\circ P_\circ$ es, evidentemente, una componente de ruido. Calculemos cuánto vale la corriente total de mezcla:

$$\bar{i}_T = R[\sqrt{2}E_{tot}]^2 = R[\sqrt{2}\sqrt{G_\circ P_\circ}\sin(\omega_1 t) + \sqrt{2}\sqrt{S_{ASE}d\omega_2}\sin(\omega_2 t + \Delta\Phi)]^2 \Rightarrow$$

$$\bar{i}_M = 4R\sqrt{G_\circ P_\circ}\sqrt{S_{ASE}d\omega_2} < \sin(\omega_1 t)\sin(\omega_2 t + \Delta\Phi) >=$$

$$= 2R\sqrt{G_\circ P_\circ}\sqrt{S_{ASE}d\omega_2} < \cos[(\omega_1 - \omega_2)t + \Delta\Phi] > \quad (11.19)$$

En la ecuación (11.19) $\Delta\Phi$ representa el desfase que hay entre la señal útil y el ruido ASE. Una cuestión que debemos dilucidar es si $\Delta\Phi$ varía rápida o lentamente en comparación con el tiempo de definición de un bit. En el primer caso, al promediar [4], obtendríamos un

[4]Suponemos que el receptor integra la señal durante el tiempo de definición del bit.

valor nulo y en el segundo el promedio daría un valor distinto de cero que tendríamos que calcular. Para averiguarlo, debemos recordar que el ruido ASE tiene un ancho de banda efectivo limitado e igual a $2B_E$. Por tanto, el tiempo de coherencia de este ruido (el tiempo durante el cual $\Delta\Phi$ no varía apreciablemente) resulta ser $t_{coh} = \frac{1}{B_R} = \frac{1}{2B_E} = T_{bit}$ [69]–[78]. Deducimos, pues, que $\Delta\Phi$ permanece prácticamente constante y, por comodidad, lo tomamos igual a cero:

$$\bar{i}_M = 2R\sqrt{G_\circ P_\circ}\sqrt{S_{ASE} d\Delta\omega}\cos(\Delta\omega t) \Rightarrow \bar{i^2}_M = 4R^2 G_\circ P_\circ \sqrt{S}_{ASE}[\tfrac{1}{2} + \tfrac{1}{2}\cos(2\Delta\omega t)]d\Delta\omega \Rightarrow$$

$$\bar{i^2}_{totM} = 4R^2 G_\circ P_\circ S_{ASE}\int_{-B_E}^{B_E}[\frac{1}{2} + \frac{1}{2}\cos(2\Delta\omega t)]d\Delta\omega = 4R^2 G_\circ P_\circ S_{ASE} B_E \qquad (11.20)$$

Ya estamos en condiciones de calcular el factor de ruido F del amplificador óptico. Sabemos que a la entrada el único ruido presente es el ruido cuántico asociado a la señal, mientra que a la salida tenemos dos términos: el ruido cuántico asociado a la señal a amplificada y el ruido de mezcla. Como son estadísticamente independientes, podemos sumar sus varianzas.

$$\left.\begin{array}{l} \left.\dfrac{S}{N}\right)_{in} = \dfrac{RP_\circ}{2qB_E} \\ \left.\dfrac{S}{N}\right)_{out} = \dfrac{RG_\circ P_\circ}{2qB_E + 4RS_{ASE}B_E} \end{array}\right\} \Rightarrow F = \frac{\dfrac{RP_\circ}{2qB_E}}{\dfrac{RG_\circ P_\circ}{2qB_E + 4RS_{ASE}B_E}} =$$

$$= \frac{2qB_E[1 + 2\dfrac{R}{q/hf}n_{sp}(G_\circ - 1)]}{2qB_E G_\circ} = \frac{1 + 2\eta n_{sp}(G_\circ - 1)}{G_\circ} \qquad (11.21)$$

En la fórmula (11.21) η es el rendimiento cuántico interno del fotodetector y dado que estamos considerando un fotodetector ideal, tomamos $\eta = 1$. Si, además, tenemos en cuenta que la ganancia del amplificador suele ser muy alta, resulta:

$$F \simeq 2n_{sp} \qquad (11.22)$$

donda n_{sp} es la tasa de emisión espontánea en exceso.

11.5. Modelo del EDFA en copropagación

En la fig. {11.7} hemos representado un amplificador óptico de fibra dopada con erbio con una configuración de bombeo en copropagación. Tal como se observa en la figura, en los extremos se han colocado dos aisladores que sirven para evitar que las reflexiones parásitas en la fibra conviertan el amplificador en un láser [5], ya que el aislador sólo permite que la luz

[5] Recordemos que un láser no es más que un medio material con ganancia que está limitado en sus extremos por dos espejos.

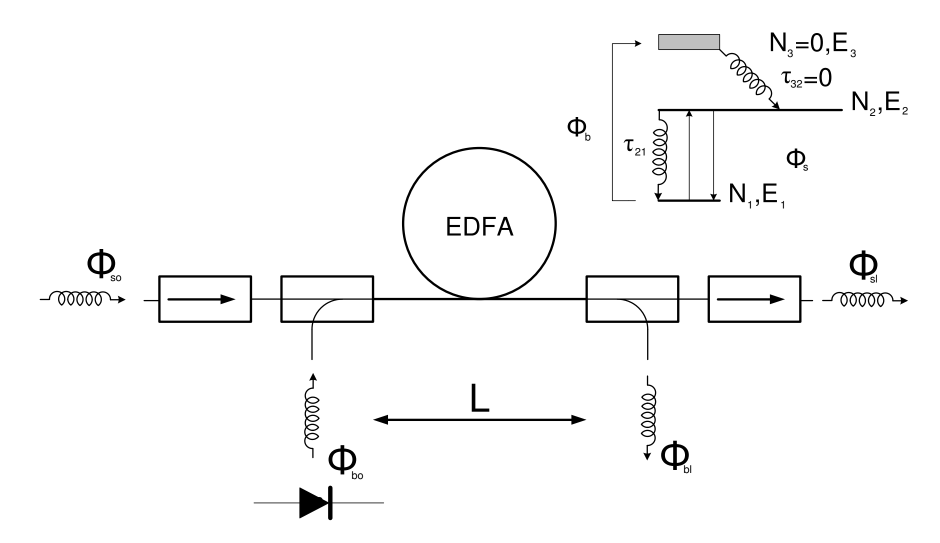


Figura 11.7: Amplificador óptico de fibra dopada con erbio. Bombeo en copropagación.

se propague en un único sentido. El multiplexor de entrada sirve para introducir en la fibra el flujo de bombeo (que puede hacerse a $\lambda = 980nm$ o bien a $\lambda = 1480nm$), mientras que el desmultiplexor de salida permite extraer el flujo de bombeo excedente, que si el EDFA está bien diseñado debe ser muy pequeño. La diferencia entre bombear a 980nm y hacerlo a 1480nm consiste en que el factor de ruido del amplificador es más bajo en el primer caso, mientras que la potencia de saturación es más alta en el segundo (tal como indicamos en la introducción de este capítulo, los modelos que vamos a deducir en este y el próximo apartado no nos van a permitir predecir este comportamiento).

En la fig. {11.7} también se representa el diagrama de estados energéticos accesibles para los iones Er^{3+} con los que se ha dopado el núcleo de la fibra. El estado "3" es, como en un láser de tres niveles, inestable, por lo que el tiempo medio de permanencia en este estado es muy pequeño: $\tau_{32} \simeq 0$ y, en consecuencia, entre los estados "0" y "3" sólo puede haber absorción y nunca emisión estimulada (ya que $N_3 \simeq 0$). El flujo de bombeo Φ_B sirve, precisamente, para estimular la absorción entre los estados "0" y "3" y conseguir que los iones caigan rápidamente al estado "2". De esta manera se crea una inversión de población entre los estados "1" y "2" y se producen más fotones por emisión estimulada que los que se roban por absorción. La constante de tiempo de la transición "1–2", τ_{21}, es lo suficientemente grande como para que resulte posible mantener la inversión de población y el número de fotones generados por emisión espontánea sea bajo.

Para simplificar, supondremos que los factores de confinamiento asociados a los flujos de bombeo y señal son iguales (lo cual no es cierto). Planteemos, a continuación, las ecuaciones diferenciales de evolución de los flujos, la conservación del número total de iones, la ecuación diferencial que describe la variación del número de iones en el estado "1" (imponemos condiciones estacionarias, ya que buscamos un modelo estático del EDFA) y las condiciones de contorno del problema:

$$\left.\begin{aligned}
\frac{d\Phi_B}{dz} &= -\Gamma\sigma_B N_0 \Phi_B \\
\frac{d\Phi_S}{dz} &= \Gamma\sigma_S (N_1 - N_0)\Phi_S \\
N_{Er^{3+}} &= N_0 + N_1 \\
\frac{dN_1}{dt} &= \Gamma\sigma_B N_\circ \Phi_B - \Gamma\sigma_S (N_1 - N_0)\Phi_S - \frac{N_1}{\tau_{10}} = 0 \\
z = 0;\quad \Phi_B &= \Phi_{B0},\quad \Phi_S = \Phi_{S0}
\end{aligned}\right\} \tag{11.23}$$

Estas ecuaciones dependen de un número excesivamente grande de parámetros, a saber: Γ, σ_S, σ_B, τ_{10}, σ_B, σ_S, $N_{Er^{3+}}$ y L. Para reducir este número vamos a *adimensionalizar* las ecuaciones, expresándolas en función de *variables adimensionales* que toman valores comprendidos entre 0 y 2:

$$\left.\begin{aligned}
\alpha_B = \frac{\Phi_B}{\Phi_{B0}};\quad \alpha_S = \frac{\Phi_S}{\Phi_{B0}} \\
\beta_0 = \frac{2N_0}{N_{Er^{3+}}};\quad \beta_1 = \frac{2N_1}{N_{Er^{3+}}} \\
\xi = \frac{z}{z_c};\quad \alpha_{B0} = 1;\quad \alpha_{S0} = \frac{\Phi_{S0}}{\Phi_{B0}}
\end{aligned}\right\} \tag{11.24}$$

El flujo de la señal a la entrada del amplificador (en $z = 0$) suele ser mucho más pequeño que el de bombeo. Además, en una situación ideal (que nunca se verifica exactamente en la práctica), en $z = L$ todos los fotones del flujo de bombeo habrían pasado al de la señal, de manera que como mucho, el flujo de señal a la salida sería igual a la suma de los flujos de bombeo y señal a la entrada, motivo por el que se toma Φ_{B0} como flujo de adimensionalización. La distancia característica de adimensionalización, z_c, no tiene por qué coincidir con L (de hecho, la mayor parte de las veces no coincide). El significado físico de z_c es la distancia a lo largo de la que los dos términos de la primera ecuación del sistema (11.23) son del mismo orden de magnitud. Expresamos a continuación el sistema (11.23) en función de las variables adimensionales definidas en (11.24):

$$\left.\begin{aligned}
\frac{\Phi_{B0}}{z_c}\frac{d\alpha_B}{d\xi} &= -\frac{\Gamma\sigma_B N_{Er^{3+}}\Phi_{B0}}{2}\beta_0\alpha_B \\
\frac{\Phi_{B0}}{z_c}\frac{d\alpha_S}{d\xi} &= \frac{\Gamma\sigma_S N_{Er^{3+}}\Phi_{B0}}{2}(\beta_1-\beta_0)\alpha_S \\
N_{Er^{3+}} &= \frac{N_{Er^{3+}}}{2}(\beta_0+\beta_1) \\
-\frac{\Gamma\sigma_B N_{Er^{3+}}\Phi_{B0}}{2}\beta_0\alpha_B + \frac{\Gamma\sigma_B N_{Er^{3+}}\Phi_{B0}}{2}(\beta_1-\beta_0)\alpha_S - \frac{N_{Er^{3+}}}{2\tau_{10}}\beta_1 &= 0 \\
\xi = 0;\quad \alpha_B = 1,\quad \alpha_S &= \frac{\Phi_{S0}}{\Phi_{B0}}
\end{aligned}\right\} \qquad (11.25)$$

Para que los dos términos de la primera ecuación del sistema (11.25) sean del mismo orden de magnitud, tomamos $z_c = \frac{2}{\Gamma\sigma_B N_{Er^{3+}}}$, con lo que queda:

$$\left.\begin{aligned}
\frac{d\alpha_B}{d\xi} &= -\beta_0\alpha_B \\
\frac{d\alpha_S}{d\xi} &= a(\beta_1-\beta_0)\alpha_S \\
\beta_0+\beta_1 &= 2 \\
\beta_0\alpha_B - a(\beta_1-\beta_0)\alpha_S - \frac{\beta_1}{b} &= 0 \\
\xi = 0;\quad \alpha_{B0} = 1,\quad \alpha_{S0} &= \frac{\Phi_{S0}}{\Phi_{B0}}
\end{aligned}\right\} \qquad (11.26)$$

Comparando los sistemas de ecuaciones (11.23) y (11.26) vemos que hemos reducido el número de parámetros a dos $a := \frac{\sigma_S}{\sigma_B}$ y $b := \tau_{10}\Gamma\sigma_B\Phi_{B0}$. En los amplificadores ópticos se cumple, aproximadamente, que $b \gg 1$ (es decir, la aportación de la emisión espontánea en la variación del número de átomos que ocupan el estado "1" es poco importante). Si ahora resolvemos en $\beta_\circ$ y β_1 el sistema formado por las ecuaciones tercera y cuarta de (11.26) queda:

$$\left.\begin{aligned}
\beta_0+\beta_1 &= 2 \\
\beta_0\alpha_B - a(\beta_1-\beta_0)\alpha_S - \frac{\beta_1}{b} &= 0
\end{aligned}\right\} \Rightarrow \left.\begin{aligned}
\beta_0 &\simeq \frac{2a\alpha_S}{\alpha_B+2a\alpha_S} \\
\beta_1 &\simeq \frac{2\alpha_B+2a\alpha_S}{\alpha_B+2a\alpha_S}
\end{aligned}\right\} \qquad (11.27)$$

Vemos que al despreciar la emisión espontánea, podemos expresar las poblaciones de los estados "0" y "1" en función de los flujos de señal y bombeo. Si ahora sustituimos (11.27) en las ecuaciones primera y segunda de (11.26) queda:

$$\left.\begin{aligned} \frac{d\alpha_B}{d\xi} &= -\frac{2a\alpha_B\alpha_S}{\alpha_B + 2a\alpha_S} \\ \frac{d\alpha_S}{d\xi} &= \frac{2a\alpha_B\alpha_S}{\alpha_B + 2a\alpha_S} \end{aligned}\right\} \tag{11.28}$$

Sumando las dos ecuaciones (11.28) comprobamos que se $\frac{d}{d\xi}(\alpha_B + \alpha_S) = 0$ ya que al despreciar el término de emisión espontánea, cada fotón perdido por el flujo de bombeo se traduce en un fotón ganado por el flujo de la señal. Si particularizamos en $\xi = 0$ queda $\alpha_B(\xi) + \alpha_S(\xi) = 1 + \frac{\Phi_{S0}}{\Phi_{B0}} \simeq 1$, que sustituido en (11.28) resulta:

$$\left.\begin{aligned} \frac{d\alpha_S}{d\xi} &= -\frac{2a\alpha_S(1-\alpha_S)}{1 + 2a\alpha_S - \alpha_S} \\ \xi = 0;\quad \alpha_S &= \alpha_{S0} = \frac{\Phi_{S0}}{\Phi_{B0}} \end{aligned}\right\} \Rightarrow \ln \frac{\left(\frac{\alpha_{SL}}{\alpha_{S0}}\right)^{\frac{1}{2a}}}{\frac{\alpha_{SL}-1}{\alpha_{S0}-1}} = \frac{L}{z} \Rightarrow G = \left(1 - \frac{\Phi_{SL}}{\Phi_{B0}}\right)^{2a} e^{\frac{2aL}{z_c}} \tag{11.29}$$

Donde $a := \dfrac{\sigma_S}{\sigma_B}$ y $z_c := \dfrac{2}{\Gamma\sigma_B N_{Er^{3+}}}$.

11.6. Modelo del EDFA en contrapropagación

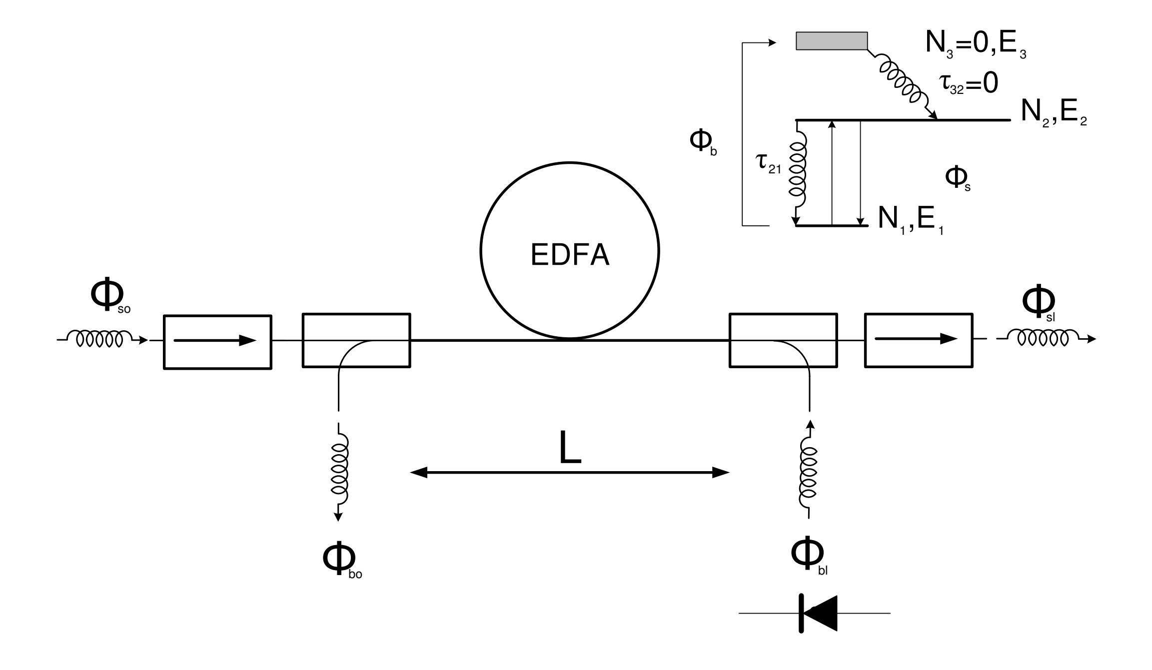


Figura 11.8: Amplificador óptico de fibra dopada con erbio. Bombeo en contrapropagación.

En la fig. {11.8} se muestra un EDFA con bombeo en contrapropagación. La diferencia fundamental con respecto al caso anterior consiste en que el bombeo se realiza desde el extremo de la fibra, con lo que la zona donde se bombea con más intensidad coincide con la región donde el flujo de la señal útil es mayor. Como consecuencia, la potencia de saturación del amplificador es más alta, aunque el modelo que vamos a deducir no es capaz de predicir este comportamiento.

$$\left.\begin{aligned}
\frac{d\Phi_B}{dz} &= \Gamma\sigma_B N_0 \Phi_B \\
\frac{d\Phi_S}{dz} &= \Gamma\sigma_S (N_1 - N_0)\Phi_S \\
N_{Er^{3+}} &= N_0 + N_1 \\
\frac{dN_1}{dt} &= \Gamma\sigma_B N_0 \Phi_B - \Gamma\sigma_S (N_1 - N_0)\Phi_S - \frac{N_1}{\tau_{10}} = 0 \\
z = 0 \;\; \Phi_S &= \Phi_{S0}; \quad z = L \;\; \Phi_B = \Phi_{BL}
\end{aligned}\right\} \tag{11.30}$$

Comparando las ecuaciones (11.30) con (11.23) advertimos de que la única diferencia con el caso anterior es que el flujo de bombeo crece con la coordenada longitudinal z, ya que el láser de bombeo está colocado al final del amplificador.

De nuevo, estas ecuaciones dependen de un número grande de parámetros: Γ, σ_S, σ_B, τ_{10}, σ_B, σ_S, $N_{Er^{3+}}$ y L. Para reducir este número adimensionalizamos las ecuaciones, expresándolas en función de las siguientes variables adimensionales:

$$\left.\begin{aligned}
\alpha_B &= \frac{\Phi_B}{\Phi_{BL}}; \quad \alpha_S = \frac{\Phi_S}{\Phi_{BL}} \\
\beta_0 &= \frac{2N_0}{N_{Er^{3+}}}; \quad \beta_1 = \frac{2N_1}{N_{Er^{3+}}} \\
\xi &= \frac{z}{z_c}; \quad \alpha_{BL} = 1; \quad \alpha_{S0} = \frac{\Phi_{S0}}{\Phi_{BL}}
\end{aligned}\right\} \tag{11.31}$$

Hemos tomado Φ_{BL} como flujo de adimensionalización. Al igual que en el apartado anterior, la distancia característica de adimensionalización z_c es la distancia a lo largo de la cual los dos términos de la primera ecuación del sistema (11.30) son del mismo orden de magnitud. Expresemos el sistema (11.30) en función de las variables adimensionales definidas en (11.31).

$$\left.\begin{aligned}
\frac{\Phi_{BL}}{z_c}\frac{d\alpha_B}{d\xi} &= \frac{\Gamma\sigma_B N_{Er^{3+}}\Phi_{BL}}{2}\beta_0\alpha_B \\
\frac{\Phi_{BL}}{z_c}\frac{d\alpha_S}{d\xi} &= \frac{\Gamma\sigma_S N_{Er^{3+}}\Phi_{BL}}{2}(\beta_1-\beta_0)\alpha_S \\
N_{Er^{3+}} &= \frac{N_{Er^{3+}}}{2}(\beta_0+\beta_1) \\
-\frac{\Gamma\sigma_B N_{Er^{3+}}\Phi_{BL}}{2}\beta_0\alpha_B + \frac{\Gamma\sigma_B N_{Er^{3+}}\Phi_{BL}}{2}&(\beta_1-\beta_0)\alpha_S - \frac{N_{Er^{3+}}}{2\tau_{10}}\beta_1 = 0 \\
\xi = \frac{L}{z_c}\quad \alpha_B = 1;\quad \xi = 0\quad \alpha_S &= \frac{\Phi_{S0}}{\Phi_{BL}}
\end{aligned}\right\} \tag{11.32}$$

Para que los dos términos de la primera ecuación del sistema (11.32) sean del mismo orden de magnitud, tomamos $z_c = \frac{2}{\Gamma\sigma_B N_{Er^{3+}}}$, con lo que queda:

$$\left.\begin{aligned}
&\frac{d\alpha_B}{d\xi} = \beta_0\alpha_B \\
&\frac{d\alpha_S}{d\xi} = a(\beta_1-\beta_0)\alpha_S \\
&\beta_0+\beta_1 = 2 \\
&\beta_0\alpha_B - a(\beta_1-\beta_0)\alpha_S - \frac{\beta_1}{b} = 0 \\
&\xi = 0;\quad \alpha_{S0} = \frac{\Phi_{S0}}{\Phi_{BL}}\quad \alpha_{BL} = 1
\end{aligned}\right\} \tag{11.33}$$

Comparando los sistemas de ecuaciones (11.23) y (11.26) vemos que hemos reducido el número de parámetros a dos $a := \frac{\sigma_S}{\sigma_B}$ y $b := \tau_{10}\Gamma\sigma_B\Phi_{BL}$. Al igual que en el caso anterior se cumple, aproximadamente, $b \gg 1$. Si ahora resolvemos en β_0 y β_1 el sistema formado por las ecuaciones tercera y cuarta de (11.33) queda:

$$\left.\begin{aligned}
&\beta_0+\beta_1 = 2 \\
&\beta_0\alpha_B - a(\beta_1-\beta_0)\alpha_S - \frac{\beta_1}{b} = 0
\end{aligned}\right\} \Rightarrow \left.\begin{aligned}
\beta_0 &\simeq \frac{2a\alpha_S}{\alpha_B+2a\alpha_S} \\
\beta_1 &\simeq \frac{2\alpha_B+2a\alpha_S}{\alpha_B+2a\alpha_S}
\end{aligned}\right\} \tag{11.34}$$

Si ahora sustituimos (11.34) en las ecuaciones primera y segunda de (11.33) queda:

$$\left.\begin{aligned} \frac{d\alpha_B}{d\xi} &= \frac{2a\alpha_B\alpha_S}{\alpha_B + 2a\alpha_S} \\ \frac{d\alpha_S}{d\xi} &= \frac{2a\alpha_B\alpha_S}{\alpha_B + 2a\alpha_S} \end{aligned}\right\} \tag{11.35}$$

Restando las ecuaciones (11.35) comprobamos que se verifica $\frac{d}{d\xi}(\alpha_B - \alpha_S) = 0$ ya que al despreciar el término de emisión espontánea, cada fotón perdido por el flujo de bombeo se traduce en un fotón ganado por el flujo de la señal. A diferencia del caso anterior, no podemos particularizar esta expresión para ningún ξ ya que no son conocidos los dos flujos simultáneamente en niguno de los dos extremos. Sin embargo, acudiendo a la definición de la ganacia del amplificador, tenemos que $\alpha_{SL} = G\alpha_{S0}$ por lo que llegamos a la integral primera $\alpha_B(\xi) - \alpha_S(\xi) = 1 - G\alpha_{S0}$. Sustituyendo en (11.35) resulta:

$$\left.\begin{aligned} &\frac{d\alpha_S}{d\xi} = \frac{2a\alpha_S(1 - G\alpha_{S0} + \alpha_S)}{2a\alpha_S + 1 - G\alpha_{S0} + \alpha_S} \\ &\xi = 0; \quad \alpha_S = \alpha_{S0} = \tfrac{\Phi_{S0}}{\Phi_{BL}} \end{aligned}\right\} \Rightarrow \ln G + 2a \ln\left(\frac{1}{1 + \alpha_{S0} - G\alpha_{S0}}\right) = \frac{2aL}{z} \Rightarrow$$

$$G \simeq \left(1 - \frac{\Phi_{SL}}{\Phi_{BL}}\right)^{2a} e^{\frac{2aL}{zc}} \tag{11.36}$$

11.7. Problemas

Problema 14: Se tiene una amplificador óptico de fibra dopada con erbio en copropagación. El coeficiente de absorción de la fibra a 980nm es de 1dB/m y los factores de confinamiento asociados a los flujos de señal y bombeo valen $\Gamma = 0,5$. El tiempo de permanencia de un ión de erbio en el estado 2 es despreciable, mientras que el asociado al estado 1 es $\tau_{10} = 50ms$. La sección cruzada de señal vale $\sigma_S = 3 \times 10^{-25} m^2$ y la de bombeo el doble. La sección efectiva asociado a ambos flujos es $A_t = 40 \mu m^2$, las potencias de bombeo y señal al principio del amplificador son $P_{B0} = 10mW$ $P_{S0} = 1\mu W$ y la longitud de onda de la señal es $\lambda_\circ = 1553nm$.
Datos adicionales: $\hbar\omega = \frac{1,24}{\lambda_\circ(\mu m)}$ y $q = 1,6 \times 10^{-19} C$.
Se pide: 1.- Calcule los valores de las constantes a, b y z_c así como el valor del flujo adimensional de la señal a la entrada del EDFA.
2.- Determine la expresión analítica de la evolución del flujo de la señal a lo largo del amplificador óptico. Calcular la longitud de la fibra de manera que el flujo adimensional de la señal a la salida sea el 95 por ciento del flujo de bombeo a la entrada.
3.- Calcule la potencia óptica de señal y bombeo a la salida del amplificador óptico.

Solución:

Apartado 1:

El coeficiente de atenuación de la fibra a 980nm expresado en unidades de m^{-1} es:

$$\alpha_{980}(m^{-1}) = \frac{\alpha_{980}(m^{-1})}{10Log_{10}e} = \frac{1}{4,343} = 0,2303m^{-1}$$

Teniendo en cuenta que $\alpha_{980} = \Gamma\sigma_B N_{Er^{3+}}$ resulta:

$$N_{Er^{3+}} = \frac{\gamma_{980}}{\Gamma\sigma_B} = \frac{0,2303}{3 \times 10^{-25}} = 7,677 \times 10^{23} m^{-3}$$

Calculemos, ahora, los parámetros a y z_c

$$z_c = \frac{2}{\Gamma\sigma_B N_{Er^{3+}}} = \frac{2}{\alpha_{980}} = 8,684m$$

$$a = \frac{\sigma_S}{\sigma_B} = \frac{1}{2}$$

Puesto que $b = \tau_{10}\Gamma\sigma_B\Phi_{B0}$, tenemos que calcular primero Φ_{B0}

$$\Phi_{B0} = \frac{P_{B0}}{\frac{1,24}{\lambda_B(\mu m)} \times q \times A_t} = \frac{10^{-2}}{\frac{1,24}{0,98} \times 1,6 \times 10^{-19} 40 \times 10^{-12}} = 1,2348 \times 10^{27} m^{-2}$$

$$b = \tau_{10}\Gamma\sigma_B\Phi_{B0} = 5 \times 10^{-2} \times 10^{-25} \times 1,2349 \times 10^{27} = 18,524$$

Calculemos también Φ_{s0}

$$\Phi_{S0} = \frac{P_{S0}}{\frac{1,24}{\lambda_S(\mu m)} \times q \times A_t} = \frac{10^{-6}}{\frac{1,24}{1,55} \times 1,6 \times 10^{-19} \times 40 \times 10^{-12}} = 1,9569 \times 10^{23} m^{-2}$$

Apartado 2:

Puesto que $b \gg 1$ y $a = \frac{1}{2}$ la ecuación diferencial de evolución del flujo de señal dentro del EDFA es:

$$\frac{d\alpha_{S\xi}}{d\xi} = \alpha_{S\xi}(\alpha_{S\xi} - 1) \Rightarrow \ln\left[\frac{\left(\frac{\alpha_{S\xi}}{\alpha_{S0}}\right)}{\left(\frac{\alpha_{S\xi}-1}{\alpha_{S0}-1}\right)}\right] = \xi \Rightarrow \frac{\alpha_{S\xi}}{\alpha_{S\xi}-1} = \frac{\alpha_{S0}}{\alpha_{S0}-1} e^{\xi}$$

Ahora lo único que hay que hacer es imponer la cndición $\alpha_{S\xi=\frac{L}{z_c}} = 0,95$:

$$\alpha_{S0} = \frac{\Phi_{S0}}{\Phi_{B0}} = \frac{1,9569 \times 10^{23}}{1,2349 \times 10^{27}} = 1,5847 \times 10^{-4}$$

$$\frac{\alpha_{0,95}}{0,95 - 1} = \frac{1,5847 \times 10^{-4}}{1,5847 \times 10^{-4} - 1} e^{\frac{L}{z_c}} \Rightarrow L = 11,69 \times z_c = 101,55m$$

Apartado 3:

Puesto que hay que calcular la potencia de la señal a la salida, calculamos primero el flujo asociado:

$$\alpha_{SL} = 0,95 \Rightarrow \Phi_{SL} = 0,95\Phi_{B0} = 1,1731 \times 10^{27} m^{-2}$$

y, después, la potencia:

$$P_{SL} = \Phi_{SL}\hbar\omega A_t = \frac{1,24}{1,55} \times 1,6 \times 10^{-19} 40 \times 10^{-12} \times 1,1731 \times 10^{27} = 6mW$$

calculemos, por último, la potencia de bombeo a la salida:

$$\Phi_{BL} = 0,05\Phi_{B0} = 6,175 \times 10^{25} m^{-2} \Rightarrow P_{BL} = \Phi_{BL}\hbar\omega A_t =$$

$$= \tfrac{1,24}{0,98} \times 1,6 \times 10^{-19} \times 10 \times 10^{-12} \times 6,174 \times 10^{25} = 0,5mW$$

Problema 15: Considere un conjunto k de amplificadores ópticos conectados en cascada, separados entre sí por fibra cuyas atenuaciones son L_1, L_2 ..., L_k.
Las ganacias de los amplificadores y sus factores de ruido son G_1 y F_1, G_2 y F_2, ..., G_k y F_k. Calcule la expresión general del factor de ruido del conjunto y optimice el diseño para los siguientes casos:
1.- Un preamplificador con un bajo nivel de ruido que ha de trabajar con señales cuyo rango dinámico es muy elevado.
2.- Amplificadores de línea que compensen la atenuación del vano de línea anterior.

Solución: Pensemos, primero, cómo afecta el factor de ruido del amplificador colocado en la primera posición al factor de ruido global. A *su entrada* el factor de ruido vale F_1 pero si consideramos que al principio del primer vano la señal útil es $\frac{1}{L_1}$ veces más fuerte, resulta que el primer amplificador contribuye al factor de ruido total de la siguiente manera:

$$F_{tot} := \frac{(S/N)_{in}}{(S/N)_{out}} = \frac{F_1}{L_1} + ...$$

La contribución del segundo amplificador también es fácil de clacular: vale F_2 en la posición que está colocado, pero a la entrada de la cadena la señal es $\frac{1}{L_1 G_1 L_2}$ veces más fuerte, por lo que su contribuición vale:

$$F_{tot} = \frac{F_1}{L_1} + \frac{F_2}{L_1 G_1 L_2} + ...$$

Razonando por inducción llegamos a:

$$F_{tot} = \frac{F_1}{L_1} + \frac{F_2}{L_1 G_1 L_2} + ... + \frac{F_k}{(\sum_{i=1}^{k-1} L_i G_i) L_k}$$

Consideremos ahora el primer caso. Evidentemente, los tramos de fibras van a ser muy cortos (latiguillos) cuyas pérdidas podemos despreciar. Necesitaremos, normalmente, sólo dos amplificadores ópticos: uno que tenga un factor de ruido lo más bajo posible (F_1), aún a costa de tener una potencia de saturación y una ganacia mediocres y un segundo que debe tener una potencia de saturación muy alta (para poder trabajar con un rango dinámico alto) aún a costa de tener una factor de ruido mediocre (F_2). El factor de ruido será:

$$F_{tot} = F_1 + \frac{F_2}{G} \simeq F_1$$

En el siguiente caso, como las ganacias de los amplificadores tienen que compensar las pérdidas, se cumplirá que $G_i L_i = 1$ con lo que tendremos:

$$F_{tot} = \frac{F_1}{L_1} + \frac{F_2}{L_2} + ... + \frac{F_k}{L_k} = \sum_{i=1}^{k} G_i F_i$$

Normalmente los amplificadores se colocan equiespaciados por los que queda:

$$F_{tot} = kGF$$

Vemos, pues, que todos los amplificadores contribuyen por igual a la figura de ruido del conjunto.

Capítulo 12

Componentes basados en el SOP

Resumen: En este capítulo vamos a estudiar dos componentes muy utilizados en comunicaciones ópticas, el aislador y el circulador óptico, cuyo funcionamiento se basa en modificar el estado de polarización de la luz. Para comprender cómo funcionan, hay que estudiar previamente una serie de elementos que modifican de manera elemental el estado de polarización de la luz (SOP), tales como los polarizadores, los retardadores de onda, rotadores (recíprocos y de Faraday) y los combinadores y separadores de haces. Para analizar estos dispositivos vamos a emplear el formalismo de Los vectores y matrices de Jones y, aunque el formalismo de las matrices de Müller y vectores de Stokes es complementario al anterior y resulta imprescindible para estudiar fenómenos tan importantes como la dispersión de polarización, no los vamos a utilizar por su complejidad matemática. Empezaremos este capítulo dando una breve introducción a los estados de polarización de la luz ya que es necesario conocer las distintas bases de descomposición de un estado de polarización arbitrario. Los aisladores y circuladores descritos en este capítulo son de tecnología micro–óptica, aunque existen productos comerciales basados en otras tecnologías, tales como la de cristales fotónicos.

12.1. Estado de polarización de la luz

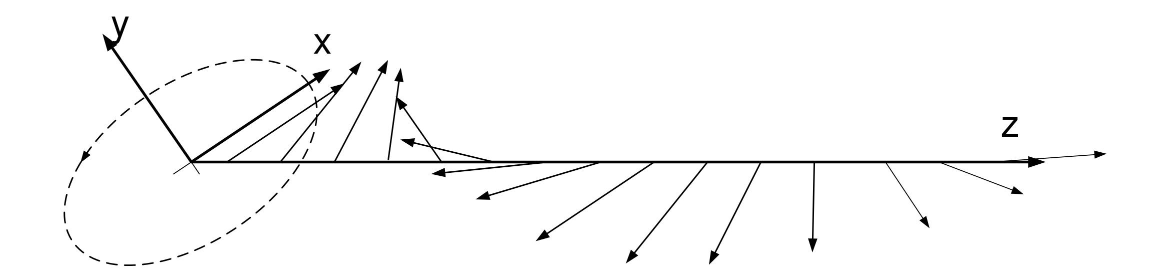


Figura 12.1: Estado de polarización (SOP) elíptico.

Consideremos un haz de luz polarizada y monocromática propagándose por un medio isótropo. Este haz luminoso lo describimos por medio de una onda localmente plana cuyo vector de onda $\vec{k}$ coincide en dirección y sentido con el vector de Poynting, mientras que las magnitudes vectoriales campo eléctrico y magnético se mantienen perpendiculares entre sí y con respecto a la dirección de propagación (es decir, la dirección a la que apunta $\vec{k}$). El extremo del campo eléctrico (o el del campo magnético) recorre, en el caso más general, una trayectoria helicoidal cuya proyección sobre el plano perpendicular a $\vec{k}$ es una elipse [17], tal como se indica en la fig. {12.1}. Diremos que la helicidad del estado de polarización es positiva si, al contemplar la onda alejándose de nosotros, el extremo del vector campo eléctrico (o magnético) recorre la hélice en el mismo sentido que las agujas del reloj y en caso contrario diremos que la helicidad es negativa. En la fig. {12.1} la helicidad es, pues, positiva.

La expresión matemática genérica del vector campo eléctrico es:

$$\vec{E} = \mathcal{R}[\vec{A}e^{j\omega_\circ(t-\frac{zn}{c_\circ})}] \tag{12.1}$$

La única restricción que imponemos $\vec{A}$ es que sea perpendicular a $\vec{k}$. Por tanto, si fijamos $z = 0$ y analizamos la evolución temporal del campo eléctrico en este plano, tenemos que siempre lo podemos descomponer en una suma de dos fasores espacialmente perpendiculares entre sí (orientados, por ejemplo, según los ejes x e y) que tienen, en general, amplitudes y referencias de fase distintas:

$$\vec{E} = \vec{A}e^{j\omega_\circ t} = a_x e^{j\Phi_x} e^{j\omega_\circ t} + a_y e^{j\Phi_y} e^{j\omega_\circ t} \tag{12.2}$$

12.2. Vectores y matrices de Jones

En el caso más general, la composición de estos dos movimientos armónicos, hace que el extremo del vector campo eléctrico (y el magnético) trace una elipse que no necesariamente está orientada según los ejes ”x” e ”y” (estos son sus ejes principales sólo en el caso en que el desfase entre ambos fasores sea igual a $\frac{\pi}{2}$ radianes). El estado de polarización (SOP) de la luz lo vamos a representar por medio de un vector columna (vector de Jones) [78] cuyas

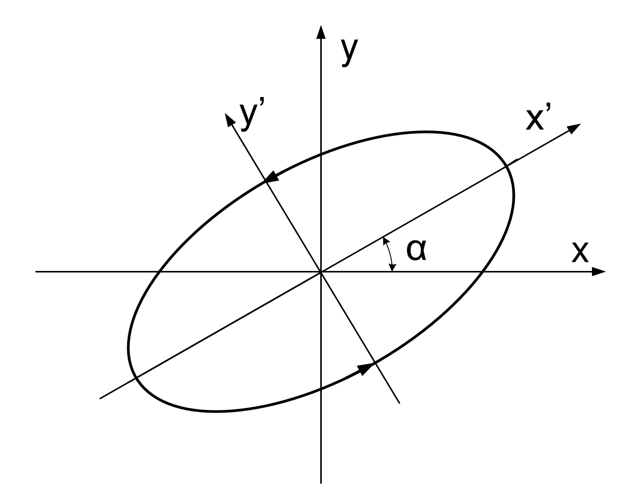


Figura 12.2: Cambio de ejes para que el fasor y' se adelante $\frac{\pi}{2}$ radianes con respecto al fasor x'. La helicidad es positiva ya que al aplicar la regla del sacacorchos, el eje z' avanza hacia el observador y, al mirar cómo se aleja la onda en este sentido, el giro de la hélice es dextrógiro.

componentes van a ser las amplitudes y referencias de fase de los dos fasores ortogonales entre sí. Sin embargo, lo que tiene relevancia física no es la fase absoluta de cada uno de los dos fasores, sino el desfase que hay entre ambos, por lo que los siguientes vectores de Jones son equivalentes entre sí:

$$J = \begin{bmatrix} a_x e^{j\Phi_x} \\ a_y e^{j\Phi_y} \end{bmatrix} \equiv \begin{bmatrix} a_x \\ a_y e^{j(\Phi_y - \Phi_x)} \end{bmatrix} \tag{12.3}$$

Si ahora hacemos el cambio de ejes indicado en la fig. {12.2} tendremos que el desfase entre los dos fasores es de $\frac{\pi}{2}$ radianes, estando el fasor orientado según y' adelantado.

$$J = \begin{bmatrix} a_x e^{j\Phi_x} \\ a_y e^{j\Phi_y} \end{bmatrix} \equiv \begin{bmatrix} a_x \\ a_y e^{j(\Phi_y - \Phi_x)} \end{bmatrix} \equiv \begin{bmatrix} a_x^{'} \\ j a_y^{'} \end{bmatrix} \tag{12.4}$$

El desfase antes mencionado implica que si observamos cómo se aleja la onda de nosotros, el sentido de giro de la elipse coincide con el de las agujas del reloj (es decir, el giro es dextrógiro). Se puede demostrar (y se deja como ejercicio para el lector) que el valor del ángulo α que define el cambio de ejes de la fig. {12.2} viene dado por la siguiente expresión [53]:

$$\tan(2\alpha) = \frac{2a_x a_y}{a_x^2 - a_y^2} \cos\Phi \tag{12.5}$$

En muchos casos sólo nos va a interesar el SOP (la norma del vector, que indica la potencia de la onda, no nos va a preocupar especialmente), con lo que podremos escribir:

$$J = \begin{bmatrix} a_x e^{j\Phi_x} \\ a_y e^{j\Phi_y} \end{bmatrix} \equiv \begin{bmatrix} a_x \\ a_y e^{j(\Phi_y - \Phi_x)} \end{bmatrix} \equiv \begin{bmatrix} 1 \\ \frac{a_y}{a_x} e^{-j\Phi} \end{bmatrix} \equiv \begin{bmatrix} a_x^{'} \\ j a_y^{'} \end{bmatrix} \equiv \begin{bmatrix} 1 \\ j \frac{a_y^{'}}{a_x^{'}} \end{bmatrix} \tag{12.6}$$

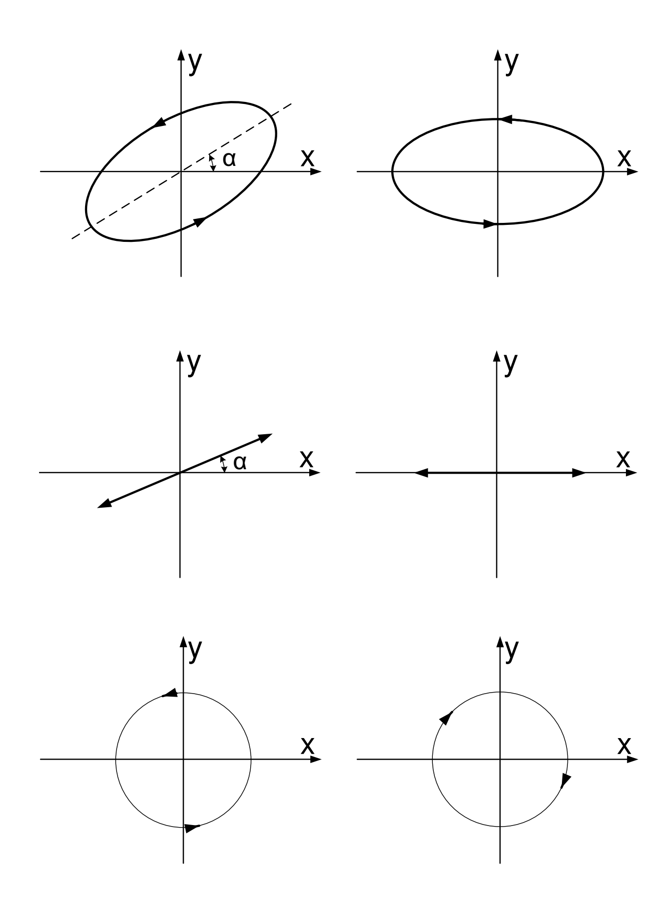


Figura 12.3: Estados de polarización particulares.

En la fig. {12.3} se muestran los estados de polarización siguientes (enumeramos de arriba hacia abajo y de izquierda a derecha);

1. SOP elíptico orientado según unos ejes no principales (helicidad positiva).
2. SOP elíptico orientado según unos ejes principales (helicidad positiva).
3. SOP lineal orientado según unos ejes no principales.
4. SOP lineal orientado según unos ejes principales.
5. SOP circular a izquierdas (helicidad positiva).
6. SOP circular a derechas (helicidad negativa).

cuyos vectores de Jones son los siguientes [78]:

$$J_1 = \begin{bmatrix} a_x \\ a_y e^{-j\Phi} \end{bmatrix} J_2 = \begin{bmatrix} a_x \\ -j a_y \end{bmatrix} J_3 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} J_4 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} J_5 = \begin{bmatrix} 1 \\ j \end{bmatrix} J_6 = \begin{bmatrix} 1 \\ -j \end{bmatrix} \quad (12.7)$$

Ahora vamos a introducir una notación [40] que va a resultar útil a lo largo de todo este capítulo [1]. Los vectores de Jones, tal como los hemos definido, los denominaremos vectores

[1]Esta es la notación de Dirac de su formulación de la mecánica cuántica [14].

del espacio directo (que, evidentemente, es un espacio bidimensional complejo) o, también, vectores "ket" y los denotaremos $|J_i^a>$, siendo i el índice que enumera las dos componentes. Los vectores fila asociados, cuyas componentes son los complejos conjugados, son vectores que pertenecen al espacio dual, los denominaremos vectores "brac" y los denotaremos $< J_i^a|$. Un vector "brac" multiplicado por un vector "ket" produce como resultado un "bracket" (corchete en inglés) que no es más que un número (complejo en general). El bracket de dos vectores de Jones es un producto interno cuyo resultado depende de qué vector tomemos como "brac" y cuál tomemos como "ket" (intercambiar los papeles de los dos vectores hace que el número resultante sea el complejo conjugado del original). El "bracket" de un estado de polarización es su norma (que, necesariamente, es un número positivo) y está relacionada con el flujo de potencia:

$$< J_i^a|J_i^a >= [A_{a1}^* \;\; A_{a2}^*] \begin{bmatrix} A_{a1} \\ A_{a2} \end{bmatrix} = A_{a1}^2 + A_{a2}^2 > 0$$

Figura 12.4: Dispositivo modificador del estado de polarización. Su representación con respecto a cualquier base de descomposición es una matriz 2×2.

Decimos que dos vectores son ortogonales si su producto interno es cero y los dos son distintos del vector nulo. Si la norma de los dos vectores ortogonales es igual a uno, diremos que los vectores son ortonormales. Cualquier par de vectores ortonormales constituyen una base adecuada sobre la que proyectar un estado de polarización abitrario, aunque lo habitual es tomar como base de referencia dos estados de polarización lineales perpendiculares entre sí y perpendiculares a la dirección de propagación. Normalmente los experimentos se realizan sobre la superficie de una mesa antivibratoria situada en un laboratorio. La dirección de propagación de la luz, normalmente se asocia al eje "z" y es paralela a la superficie de la mesa. El eje "x" (perpendicular a "z", claro está) es también paralelo a esta superficie, mientras que el eje "y" es perpendicular a la mesa. Los ejes "x", "y" y "z" forman una triedro a derechas ($dV = \widetilde{dx} \wedge \widetilde{dy} \wedge \widetilde{dz}$).

Otra base de descomposición habitual (que sirve para estudiar dispositivos con efecto giromagnético) está formada por los vectores de norma unidad asociados a los estados de

polarización circulares a derecha e izquierda. Si expresamos estos vectores con respecto a la base formada por los estados de polarización lineales orientados según los ejes ”x” e ”y” nos resulta:

$$|J_i^d>= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ j \end{bmatrix}; \quad |J_i^l>= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -j \end{bmatrix} \Rightarrow < J_i^d|J_i^l>= \frac{1}{2}[1+(-j)\times(-j)] = 0 \qquad (12.8)$$

En la fig. {12.4} se ha representado un dispositivo modificador del SOP. Al actuar sobre la onda de entrada $|J^{in}>$, produce un estado de polarización a la salida $|J^o>$. El comportamiento del dispositivo se describe por medio de una matriz cuadrada 2×2 [53] que denotaremos por medio de $\overline{\overline{T}}_{ij}$ con $i, j = 1, 2$ y que al multiplicar al vector $|J_j^{in}>$ da como resultado el vector columna $|J_i^o>= \overline{\overline{T}}_{ij}|J_j^{in}>$, o, escrito de una manera abreviada, $|J^o>= \overline{\overline{T}}|J^{in}>$, que describe el SOP de la luz a la salida. Si en vez de haber un único dispositivo hay una sucesión de ellos $\overline{\overline{T}}_{ij}^1$ (el primero que actúa sobre el haz de entrada), $\overline{\overline{T}}_{ij}^2$ (el segundo), $\overline{\overline{T}}_{ij}^3$ (el tercero) el estado de polarización resultante viene descrito por:

$$|J^o>= \{\overline{\overline{T}}_{ij}^3[\overline{\overline{T}}_{ij}^2(\overline{\overline{T}}_{ij}^1|J^{in}>)]\} \qquad (12.9)$$

donde es muy importante efectuar los productos en el orden indicado. En general las matrices que describen el comportamiento de los dispositivos modificadores del SOP no tienen por qué ser no singulares y, por tanto, el producto de éstas no disfruta de la propiedad asociativa.

12.3. Polarizadores

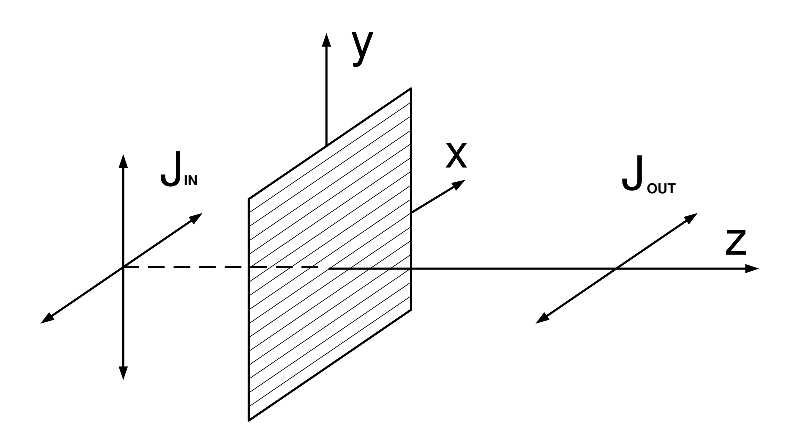


Figura 12.5: Polarizador orientado según el eje ”x”.

Los polarizadores son normalmente láminas dicroicas [53]–[55], es decir, son dispositivos fabricados con un material que presenta coeficientes de atenuación muy distintos (uno grande y otro pequeño) según dos direcciones perpendiculares. Idealmente dejan pasar sin

atenuación alguna la componente de luz linealmente polarizada según el eje privilegiado y absorben toda la luz correspondiente a la luz linealmente polarizada según la dirección perpendicular. Supongamos que el eje de baja atenuación del polarizador está alineado según la dirección "x". Si la base de descomposición son los estados de polarización lineales según los ejes "x" e "y" la matriz que describe el polarizador es:

$$\overline{\overline{T}}_x^{pol} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{12.10}$$

El comportamiento del polarizador se ha representado en la fig. {12.5}. La cuestión que podemos plantearnos a continuación es qué expresión tendría la matriz si los ejes "x" e "y" estuviese girados un ángulo arbitrario θ con respecto a los del laboratorio [104]. Esto es fácil de averiguar. En primer lugar deberíamos referir el estado de polarización a la entrada, no con respecto a los ejes del laboratorio, sino con respecto a los ejes girados asociados al polarizador. Con respecto a esos ejes, la matriz del polarizador viene dada por (12.10), por lo que tendríamos que multiplicar esta matriz por la que representa el giro (o cambio de base) antes mencionado. La matriz producto serviría para calcular el estado de polarización de la onda a la salida con respecto a los ejes girados del polarizador. Como queremos referir el estado de polarización a los ejes del laboratorio, tenemos que multiplicar la matriz anterior por la matriz asociada a un giro de ejes de valor $-\theta$, con lo que resulta:

$$\overline{\overline{T}}_\theta^{pol} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \right] = \begin{pmatrix} \cos^2\theta & \dfrac{\sin 2\theta}{2} \\ \dfrac{\sin 2\theta}{2} & \cos^2\theta \end{pmatrix} \tag{12.11}$$

La matriz correspondiente al polarizador girado un ángulo θ con respecto a los ejes del laboratorio es también singular. Esto es así porque el dispositivo absorbe toda la potencia de la onda linealmente polarizada según el eje girado θ grados con respecto al "x". De hecho, si calculamos el determianante de la matriz (12.11) comprobamos que vale cero.

12.4. Retardadores de onda

El retardador de onda [106] es un dispositivo que tiene dos ejes privilegiados y ortogonales entre sí llamados eje rápido y eje lento. Si referimos el estado de polarización de la luz entrante a la base constituida por las dos polarizaciones lineales alineadas según estos dos ejes, el dispositivo introduce un *retraso* de fase en la componente lenta con respecto a la rápida de valor Γ (Γ es el parámetro que caracteriza al retardador). En la fig. {12.6} se representa un retardador de onda cuyo eje lento está orientado segón el eje "x". La matriz de Jones asociada a este dispositivo es:

$$\overline{\overline{T}}_x^{ret} = \begin{bmatrix} e^{-j\Gamma} & 0 \\ 0 & 1 \end{bmatrix} \tag{12.12}$$

Si tomamos $\Gamma = \pi$ resulta que el retardador gira espacialmente $\frac{\pi}{2}$ radianes un haz de luz con polarización lineal a $\frac{\pi}{4}$, de manera que a la salida se obtiene otro haz de luz linealmente

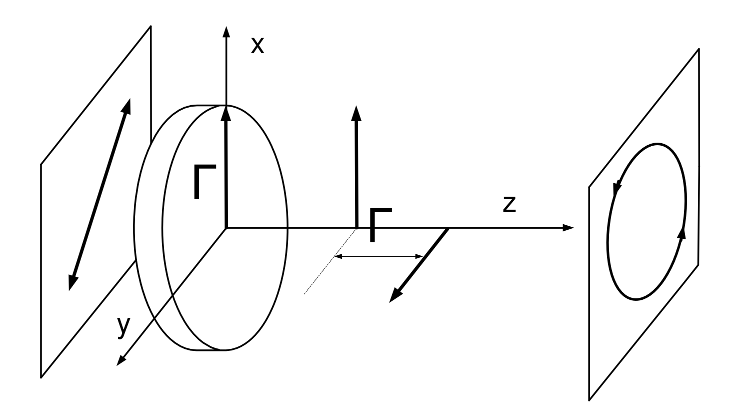


Figura 12.6: Retardador de onda cuyo eje lento está orientado según el eje "x".

polarizado a $-\frac{\pi}{4}$ y viceversa. Si tomamos como referecia la relación (12.12) y seguimos el mismo razonamiento del apartado anterior, podemos determinar la matriz asociada a un retardador cuyo eje lento forma un ángulo θ con respecto al eje "x":

$$\begin{aligned} \overline{\overline{T}}_{\theta}^{\Gamma} &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e^{-j\Gamma} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \\ &= \begin{bmatrix} \sin^2\theta + \cos^2\theta e^{-j\Gamma} & \sin\theta\cos\theta(1-e^{-j\Gamma}) \\ \sin\theta\cos\theta(1-e^{-j\Gamma}) & \cos^2\theta + \sin^2\theta e^{-j\Gamma} \end{bmatrix} \end{aligned} \quad (12.13)$$

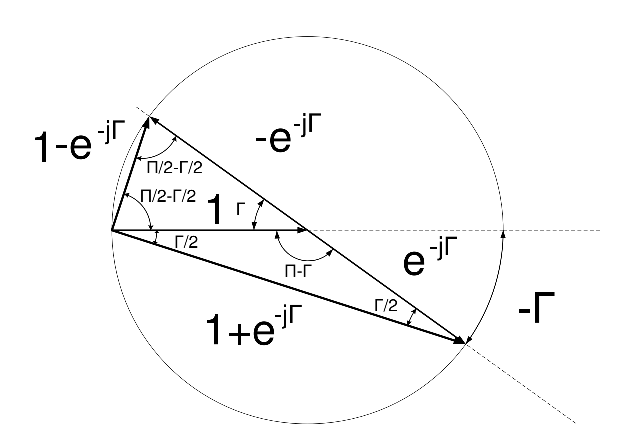


Figura 12.7: Diagrama fasorial.

Si ahora particularizamos para el valor $\theta = \frac{\pi}{4}$ queda:

$$\overline{\overline{T}}_{\frac{\pi}{4}}^{\Gamma} = \frac{1}{2} \begin{bmatrix} 1+e^{-j\Gamma} & 1-e^{-j\Gamma} \\ 1-e^{-j\Gamma} & 1+e^{-j\Gamma} \end{bmatrix} \quad (12.14)$$

Como la expresión analítica de la matriz (12.14) es difícil de interpretar vamos a representar los fasores 1, $e^{-j\Gamma}$, $1+e^{-j\Gamma}$ y $1-e^{-j\Gamma}$ en el plano. Estos fasores se han dibujado

en la fig. {12.7} y demuestran que los fasores $1 + e^{-j\Gamma}$ y $1 - e^{-j\Gamma}$ son perpendiculares entre sí estando el primero retrasado con respecto al segundo. Como además resulta evidente que $|1 + e^{-j\Gamma}| = 2\cos(\frac{\Gamma}{2})$ y $|1 - e^{-j\Gamma}| = 2\sin(\frac{\Gamma}{2})$ resulta que la matriz (12.14) es equivalente a:

$$\overline{\overline{T}}^{\Gamma}_{\frac{\pi}{4}} = \begin{bmatrix} \cos(\frac{\Gamma}{2}) & j\sin(\frac{\Gamma}{2}) \\ j\sin(\frac{\Gamma}{2}) & \cos(\frac{\Gamma}{2}) \end{bmatrix} \tag{12.15}$$

Figura 12.8: Principio físico de funcionamiento de un retardador de onda.

Si en (12.15) tomamos $\Gamma = \pi$ vemos que el dispositivo resultante transforma una onda linealmente polarizada según el eje "x" en otra, también linealmente polarizada, pero según el eje "y" y viceversa.

Los retardadores de onda se fabrican a partir de medios *uniaxiales*. Recordemos que muchos sólidos cristalinos son materiales *anisótropos*, es decir, sus propiedades físicas depende de la dirección en la que se midan. Una de esas propiedades puede ser el índice de refracción, que en medios anisótropos se describe por medio de un tensor de segundo orden de dimensiones 3×3. En estos materiales los vectores desplazamiento y campo eléctrico no están, en general, alineados. Para simplificar el análisis, vamos a suponer que el sistema de referencia del laboratorio coinciden con los ejes principales del llamado *elipsoide de índices* [53]–[106]. Con respecto a este sistema de referencia, la *matriz de índices* es diagonal, con lo que necesitamos conocer sólo tres índices de refracción distintos n_x, n_y y n_z para describir la propagación de ona onda luminosa. Un medio ópticamente anisótropo recibe el nombre de *medio biaxial* si los tres índice son distintos y el nombre de *medio uniaxial* si dos índices son iguales (índices ordinarios) y otro distinto (índice extraordinario) [53]–[106]. Normalmente el índice ordinario es menor que el extraordinario.

Si lanzamos un haz de luz polarizada por un medio uniaxial en la dirección principal "z", la componente linealmente polarizada según el eje ordinario se adelanta en fase con respecto a la componente linealmente polarizada según el eje extraordinario, tal y como se muestra en la fig. {12.8}. El valor del desfase es $\Gamma + p2\pi = k_\circ d(n_e - n_\circ) = \frac{2\pi}{\lambda_\circ} d(n_e - n_\circ)$, donde p es un número entero. De esta manera, para un valor de Γ deseado, podemos tomar una longitud d de un valor razonable (y ahí es donde entra el tomar un número p adecuado), dado que conocemos los valores $n_\circ$ y n_e de nuestro material cristalino.

12.5. Rotador de Faraday

El efecto Faraday [53]–[87] es un fenómeno que se produce en medios materiales transparentes con simetría cilíndrica en los que aparece una dirección privilegiada. El motivo es la presencia de un campo manético $\vec{H}$ alineado según dicha dirección. Para estudiar la propagación de la luz en estos medios, resulta conveniente tomar como base de descomposición los estados de polarización circulares a izquierda y derecha, ya que estos estados de polarización son autovectores con respecto a la propagación a través del medio (un estado de polarización circular, sea a derechas o izquierdas, lo sigue siendo después de atravesar un medio magnetoóptico). Se puede demostrar que los índices de refracción vistos por estos dos estados de polarización son distintos (los autovalores son distintos) y los denotaremos n_d y n_i. Por tanto, aparece un retardo de fase cuyo valor es $\Gamma + p2\pi = \frac{2\pi}{\lambda_o}(n_d - n_i)d$. Dado que la base formada por los estados de polarización circulares es poco intuitiva, vamos a ver qué ocurre con dos ondas cuyos estados de polarización son lineales y están alineados según los ejes "x" e "y" del laboratorio cuando atraviesan este medio:

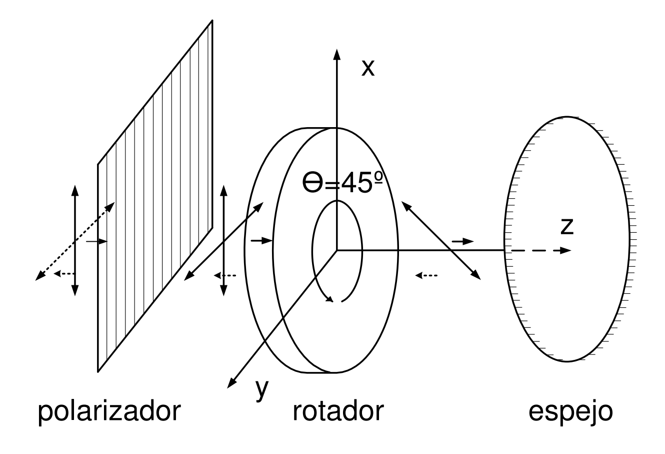


Figura 12.9: Aplicación de un rotador de Faraday.

$$
\begin{aligned}
J_x^i &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ j \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 \\ -j \end{bmatrix} \\
J_y^i &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{2j}\begin{bmatrix} 1 \\ j \end{bmatrix} - \frac{1}{2j}\begin{bmatrix} 1 \\ -j \end{bmatrix}
\end{aligned} \tag{12.16}
$$

Si ahora propagamos estos dos haces luminosos por el medio magnetoóptico nos resulta, a la salida:

$$
\begin{aligned}
J_x^o &= \frac{1}{2}e^{-j\frac{2\pi}{\lambda_o}n_d d}\begin{bmatrix} 1 \\ j \end{bmatrix} + \frac{1}{2}e^{-j\frac{2\pi}{\lambda_o}n_i d}\begin{bmatrix} 1 \\ -j \end{bmatrix} = e^{-j\frac{\pi}{\lambda_o}(n_d+n_i)d}\begin{bmatrix} \cos(\frac{\Delta\phi}{2}) \\ \sin(\frac{\Delta\phi}{2}) \end{bmatrix} \\
J_y^i &= \frac{1}{2j}e^{-j\frac{2\pi}{\lambda_o}n_d d}\begin{bmatrix} 1 \\ j \end{bmatrix} - \frac{1}{2j}e^{-j\frac{2\pi}{\lambda_o}n_i d}\begin{bmatrix} 1 \\ -j \end{bmatrix} = e^{-j\frac{\pi}{\lambda_o}(n_d+n_i)d}\begin{bmatrix} \sin(\frac{\Delta\phi}{2}) \\ \cos(\frac{\Delta\phi}{2}) \end{bmatrix}
\end{aligned} \tag{12.17}
$$

Donde $\frac{\Delta\phi}{2} + 2p\pi = \frac{\pi}{\lambda_o}(n_d - n_i)d$. Vemos, pues, que en un medio donde existe efecto Faraday, un estado de polarización lineal que se propaga por su seno sigue siendo lineal pero

gira *espacialmente* un ángulo que podemos controlar fijando el parámetro d . La relación que hay entre el ángulo de giro y el espesor del rotador es $\theta + 2p\pi = \frac{\pi}{\lambda_\circ}(n_d - n_i)d$. Esta rotación del plano de polarización fue descubierta por Faraday en 1845. Hay que destacar que se trata de un fenómeno no recíproco, es decir, si detrás del rotador de Faraday colocásemos un espejo, el estado de polarización de la onda reflejada, al pasar por el medio dos veces, una vez en un sentido y otra en el contrario, sería lineal pero girado en el espacio un ángulo igual a 2θ. La matriz de Jones de un rotador de Faraday es:

$$\overline{\overline{T}}^{\theta} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \tag{12.18}$$

En la fig. {12.9} se muestra una aplicación del rotador de Faraday (un aislador dependiente del estado de polarización). El funcionamiento es sencillo: el espejo representado no debe entenderse como tal, sino como un dispositivo que tiene un cierto coeficiente de reflexión. La hipotética onda reflejada ha pasado dos veces por el rotador de Faraday, acumulando un giro espacial de valor π, motivo por el que no puede atravesar el polarizador colocado a la entrada.

Existen también medios materiales que presentan efecto giromagnético (birrefringencia circular [53]) de manera natural y que se pueden utilizar para fabricar rotadores [87]. A diferencia de los rotadores de Faraday, los rotadores fabricados con estos materiales son *recíprocos*, por lo que no sirven para fabricar aisladores pero son indispensables en la fabricación de circuladores ópticos.

12.6. Separadores y combinadores de haz

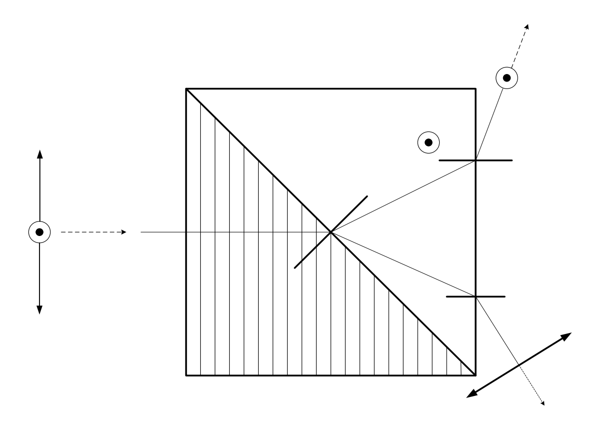


Figura 12.10: Prisma de Wollaston.

En tecnología microóptica estos dispositivos se fabrican a partir de cubos de materiales uniaxiales. El cubo tiene el eje extraordinario orientado de destintas maneras a ambos lados de un plano que bisecciona el cubo [78], tal y como se muestra en la fig. {12.10}. En la superficie de entrada (separación entre el aire y el cubo) no hay ningún cambio de dirección

en la propagación de las dos componentes linealmente polarizadas, ya que la incidencia es normal. Dentro de la primera región del cubo, la componente con polarización vertical "ve" el índice extraordinario (siempre indicamos en el cubo explícitamente la orientación del eje extraordinario) mientras que la componente con polarización horizontal, al ser perpendicular al eje extraordinario, "ve" el índice ordinario. Al llegar al plano que bisecciona el cubo, la componente vertical pasa de vibrar paralelamente al eje extraordinario, a hacerlo perpendicularmentea éste, por lo que pasa a un medio ópticamente menos denso. En consecuencia cambia su dirección de propagación, alejándose de la normal al plano bisector. Con la componente horizontal ocurre lo contrario, por lo que también cambia su dirección de propagación, esta vez acercándose a la normal al plano bisector. De esta manera conseguimos separar el haz luminoso de entrada en las dos componentes de polarización lineal (horizontal y vertical). De izquierda a derecha, tenemos un separador de haces pero si invertimos los sentidos de propagación tendremos un combinador de haces.

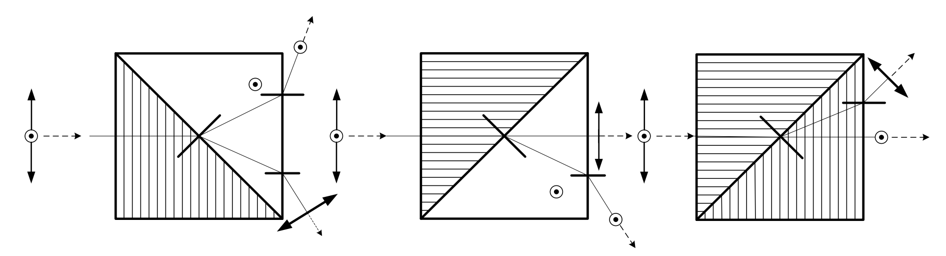


Figura 12.11: Prismas separadores de haz.

Evidentemente la configuración mostrada en la fig. {12.10} no es la única manera de separar (o combinar) un haz en sus dos componentes linealmente polarizadas. En la fig. {12.11} se muestran, junto al prisma de Wollaston, otras dos configuraciones [78].

12.7. Aisladores ópticos

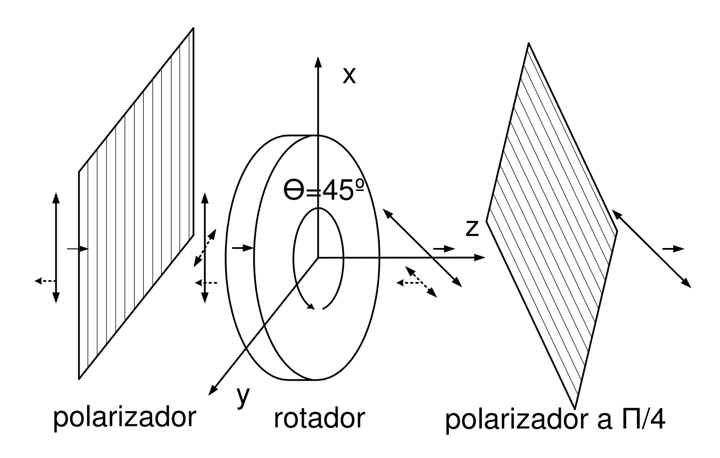


Figura 12.12: Aislador dependiente del SOP.

Este es el primero de los dos dispositivos que queríamos estudiar. Este componente permite que la luz se propague en un sentido, pero impide que lo haga en el contrario. En

los EDFAs se colocan aisladores en ambos extremos del aislador para evitar que se produzca una realimentación positiva en el amplificador y se comporte como un láser. Primero vamos a ver cómo se construye un aislador cuyo estado de polarización de entrada es conocido y, después, cómo a partir de éste, se puede fabricar uno para un estado de polarización de entrada arbitrario [84].

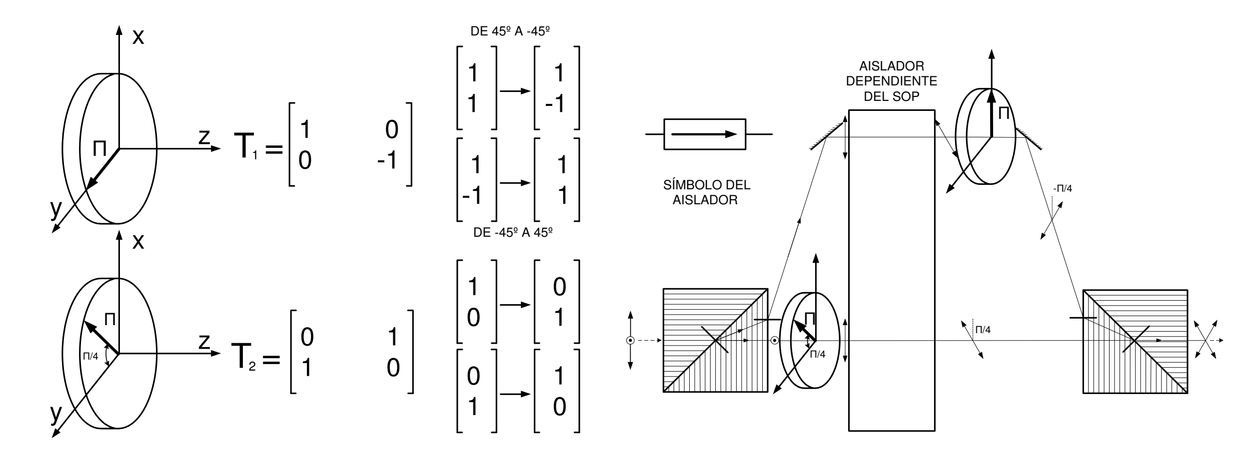


Figura 12.13: Funcionamiento de dos retardadores de onda (izquierda). Aislador independiente del estado de polarización (derecha).

En la fig. {12.12} se muestra cómo funciona este dispositivo. Si a la salida del polarizador a $\frac{\pi}{4}$ hay un componente que introduce un pequeño coeficiente de reflexión, el rotador de Faraday hace que la luz reflejada llegue al polarizador de entrada, perpendicularmente al eje "x", por lo que no puede atravesarlo.

Para comprender el funcionamiento de un aislador independiente del SOP de la luz de entrada debemos recordar, primero, el funcionamiento de los retardadores de onda. Para el caso particular de $\Gamma = \pi$ y estando orientado según el eje "y" el retardador de onda cambia la polarización lineal de entrada de $-\frac{\pi}{4}$ a $\frac{\pi}{4}$ y viceversa. Si el eje lento está orientado a $\frac{\pi}{4}$ con respecto al eje "x" hace que si la polarización lineal de entrada está alineada con el eje "x" pase a estarlo con el eje "y" y viceversa, tal y como se muestra en la fig. {12.13}. Con estos elementos y el aislador dependiente del estado de polarización podemos construir un aislador independient del SOP de la luz de entrada. Lo único que hay que hacer es, primero, separar el haz luminoso de entrada en sus dos componentes linealmente polarizadas con un prisma de Sernamont. La componente de polarización horizontal la transformamos, mediante un retardador de onda, en una onda de polarización vertical. A continuación hacemos pasar por el aislador dependiente del SOP ambas componentes, con lo que a la salida tenemos dos ondas linealmente polarizadas a $\frac{\pi}{4}$. Hacemos que una de ellas atraviese otro retardador para transformarla en una onda de polarización lineal a $\frac{-\pi}{4}$. Por último combinamos ambos haces con otro prisma de Sernamont (girado espacialmente un ángulo $\frac{\pi}{4}$). El aislador resultante se muestra en la fig. {12.13}.

12.8. Circuladores ópticos

Estos componentes son fundamentales en redes con multiplexación en longitud de onda (WDM) y su funcionamiento es idéntico a los circuladores que se utilizan en microondas: son dispositivos con tres o cuatro puertos, cuya función es sacar por el puerto siguente, y sólo por

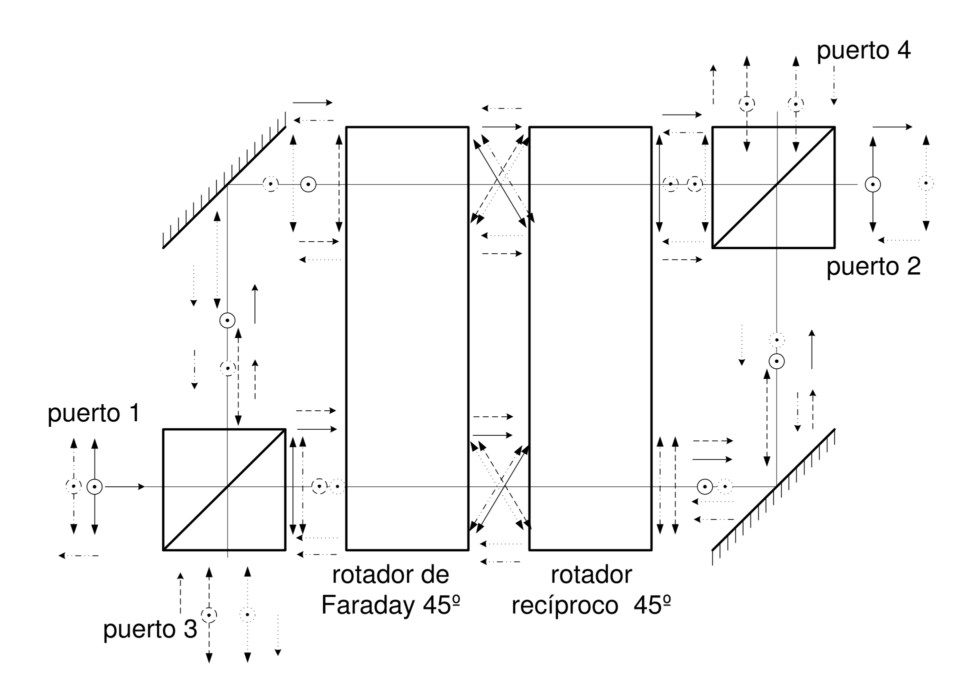


Figura 12.14: Circulador óptico.

el puerto siguiente, la señal de entrada [84]. En la fig. {12.14} se muestra la implementación en tecnología microóptica de un circulador óptico de cuatro puertos. El símbolo que aparenta ser un prisma es, en realidad, un dispositivo que deja pasar la componente de polarización vertical, reflejando la horizontal y cuyo principio de funcionamiento se basa en el denominado "ángulo de Brewster" [2]. El conjunto formado por un rotador de Faraday de $\frac{\pi}{4}$ y una rotador recíproco también de $\frac{\pi}{4}$ hace que, si una onda atraviesa los dispositivos en un sentido (de izquierda a derecha en la figura), el plano de polarización gire un ángulo de $\frac{\pi}{2}$, pero si lo atraviesa en sentido contrario no modifica el estado de polarización. De esta manera se consigue el comportamiento no recíoproco típico de un circulador. Además de la microóptica, existen otras tecnologías de fabricación de circuladores ópticos [99].

[2]En una superficie de separación entre dos dieléctricos, la luz con polarización TE (campo eléctrico vibrando paralelamente a la superficie de separación) se ve totalmente reflejada si el ángulo de incidencia toma una valor determinado, llamado ángulo de Brewster.

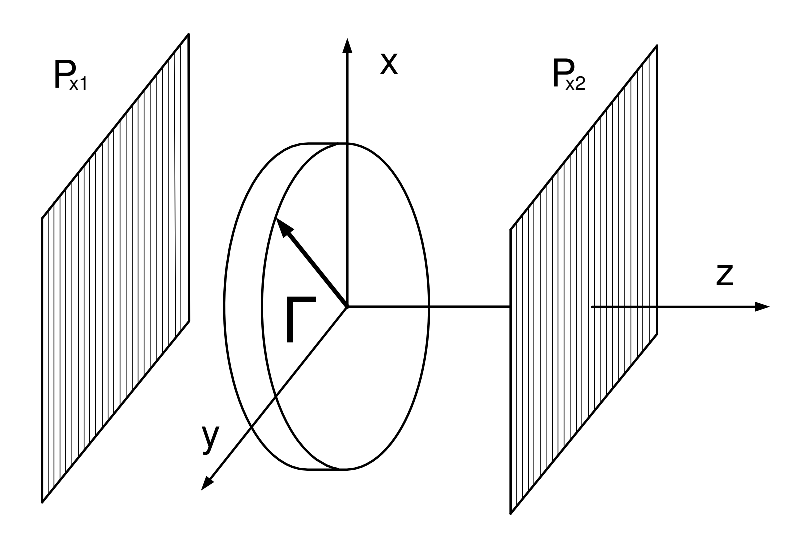


Figura 12.15: Analizador óptico.

Problema 16: Dado el analizador representado en la fig. {12.15}, se pide deducir la intensidad luminosa a la salida en función de la longitud de onda de la luz incidente.

Solución: Puesto que el eje lento del retardador de onda forma un ángulo de $\frac{\pi}{4}$ radianes con respecto a los ejes principales, la matriz de Jones asociada es:

$$\overline{\overline{T}}^{\Gamma} = \begin{bmatrix} \cos(\frac{\Gamma}{2}) & j\sin(\frac{\Gamma}{2}) \\ j\sin(\frac{\Gamma}{2}) & \cos(\frac{\Gamma}{2}) \end{bmatrix} \tag{12.19}$$

Por lo que el analizador viene descrito por el siguiente producto matricial:

$$\overline{\overline{T}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \left[\begin{pmatrix} \cos(\frac{\Gamma}{2}) & j\sin(\frac{\Gamma}{2}) \\ j\sin(\frac{\Gamma}{2}) & \cos(\frac{\Gamma}{2}) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right] = \begin{pmatrix} \cos\frac{\Gamma}{2} & 0 \\ 0 & 0 \end{pmatrix} \tag{12.20}$$

La intensidad luminosa a la salida del analizador es:

$$I = \cos^2(\frac{\Gamma}{2}) = \cos^2[\frac{\pi(n_e - n_\circ)d}{\lambda}] \tag{12.21}$$

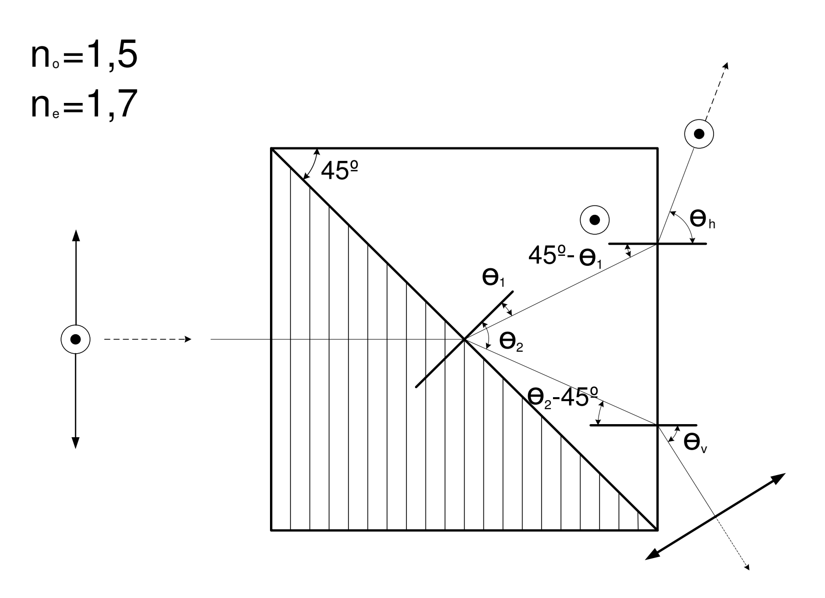


Figura 12.16: Prisma separador de Wollaston.

Problema 17: Dado el prisma separador de haces de Wollaston mostrado en la fig. {12.16} determine las direcciones de salida de las componentes de polarización horizontal y vertical.

Solución:

$$\left.\begin{array}{l} 1,7 \times \sin(\frac{\pi}{4}) = 1,5 \times \sin(\theta_2) \\ 1,5 \times \sin(\frac{\pi}{4}) = 1,7 \times \sin(\theta_1) \end{array}\right\} \Rightarrow \left.\begin{array}{l} \theta_2 = 53,263^\circ \\ \theta_1 = 38,603^\circ \end{array}\right\} \tag{12.22}$$

$$\left.\begin{array}{l} 1,7 \times \sin(\frac{\pi}{4} - \theta_1) = 1 \times \sin(\theta_h) \\ 1,5 \times \sin(\theta_2 - \frac{\pi}{4}) = 1 \times \sin(\theta_v) \end{array}\right\} \Rightarrow \left.\begin{array}{l} \theta_h = 10,918^\circ \\ \theta_v = 12,449^\circ \end{array}\right\} \tag{12.23}$$

Capítulo 13

Filtros ópticos

Resumen: En este capítulo vamos a estudiar los filtros ópticos pasobanda que se emplean en sistemas DWDM. Estos filtros pueden ser fijos, en cuyo caso se fabrican en tecnología todo fibra, o bien sintonizables, en cuyo caso se fabrican en tecnología microóptica. Los primeros reciben el nombre de fibras de difracción de Bragg (FBGs) y se emplean tanto en la selección de canales como en la fabricación de compensadores pasivos de la dispersión. Los segundos son cavidades pasivas Fabry–Pérot y se utilizan, también, para la selección de canales en aquellos casos en los que el nodo de la red deba tener la posibilidad de sintonizar distintas longitudes de onda (distintos canales). El estudio de los filtros se va a basar en el formalismo matricial de bipuertas, más concretamente, en la descripción de bipuertas mediante matrices de transmisión.

13.1. Introducción

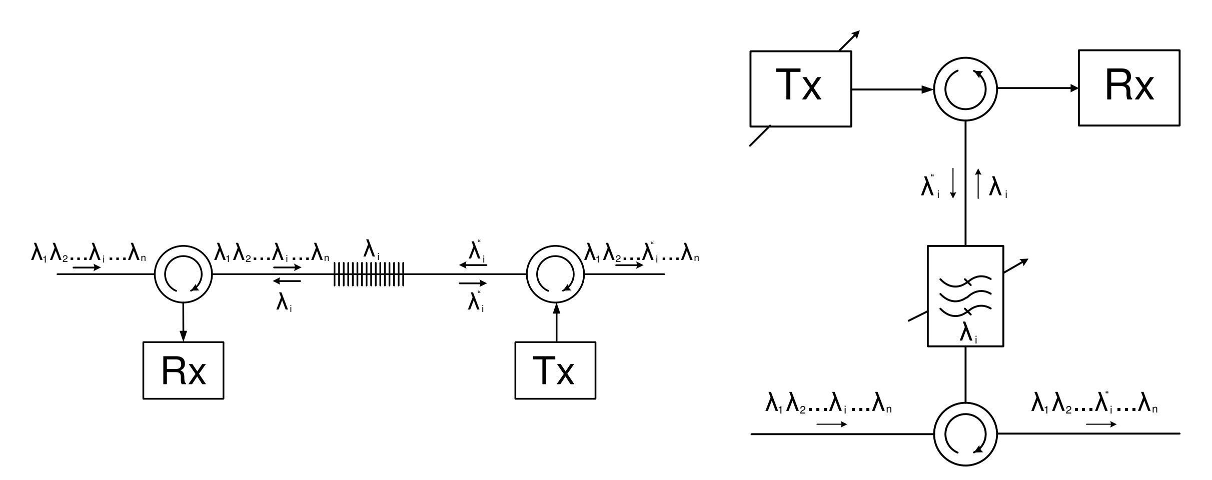


Figura 13.1: Multiplexor "add & drop" fijo (izquierda) y sintonizable (derecha).

Los filtros ópticos paso–banda tienen una importancia primordial en los sistemas DWDM. En primer lugar sirven para la implementación física de los denominados multiplexores "add & drop" [8]–[84], cuya misión es extraer la información del canal "i", que está soportada por la longitud de onda λ_i, y la sustituirla por otra distinta codificada en la misma longitud de onda, que denotamos mediante λ_i' para indicar que la información es distinta a la extraida. Estos multiplexores se construyen siempre a partir de filtros pasobanda (fijos o sintonizables según el tipo de multiplexor) y circuladores ópticos, tal y como se observa en la fig. {13.1}.

Si el multiplexor funciona a una longitud de onda fija, el filtro se suele implementar mediante una fibra de difracción de Bragg (FBG) [60], que no es más que trozo de fibra monomodo en la que hay una perturbación longitudinal del índice de refracción del núcleo. Esta perturbación periódica (normalmente senoidal) refleja [1] una determinada longitud de onda (la longitud de onda de Bragg) y transmite el resto. Los multiplexores "add & drop" fijos se colocan siempre en serie con la fibra óptica por la que se propagan los canales multiplexados en longitud de onda, tal y como se muestra en la parte izquierda de la fig. {13.1}. Este filtro no es sintonizable, ya que no se puede modificar el periodo de la perturbación introducida en la FBG.

Si el multiplexor tiene que ser sintonizable, entonces el filtro se implementa físicamente mediante una cavidad Fabry–Pérot que funciona en transmisión, es decir, deja pasar una determinada longitud de onda, reflejando el resto. Los filtros de cavidad Fabry–Pérot son sintonizables ya que en ellos se puede modificar fácilmente la longitud de la cavidad y, por tanto, la longitud de onda transmitida. Los multiplexores "add & drop" se coloca en derivación, tal y como se mueastra en la parte de la derecha de la fig. {13.1}.

Otra aplicación importante de los filtros en sistemas DWDM es la realización física de compensadores pasivos de la dispersión [60]. Estos filtros, que se fabrican en tecnología

[1] La inmensa mayoría de estos filtros funcionan en reflexión y, muy raramente lo hacen en transmisión.

todo fibra (son FBGs), tienen una variación lineal del periodo de Bragg, de manera que las distintas longitudes de onda se reflejan en distintos puntos dentro de la FBG, con lo que se introduce un retardo de grupo que depende de la longitud de onda y que puede, por tanto, compensar así la dispersión cromática. Es posible, pues, construir regeneradores "2R" en el dominio óptico (compensación de la atenuación con amplificadores ópticos y de la dispersión con redes de difracción de Bragg). Lo único que falta para poder implementar un regenerador "3R" en el dominio óptico es la "resincronización" de los pulsos con el reloj de transmisión (aunque es un campo donde se está haciendo una intensa labor de investigación, siendo los dispositivos basados en SOAs los más firmes candidatos para realizar esta función [25]).

13.2. Filtro de cavidad Fabry–Pérot.

Los filtros de cavidad Fabry–Pérot los estudiamos en el capítulo ocho donde, utilizando el formalismo de matrices de scattering y transmisión, llegamos a la conclusión de que la función de transferencia en transmisión de estos filtros viene descrita por la siguiente expresión:

$$T(f) := |S_{21}|^2 = \frac{(1-R)^2}{(1-R)^2 + 4R\sin^2(\frac{\pi f}{FSR})} \quad FSR := \frac{c_\circ}{2n_2 L} \tag{13.1}$$

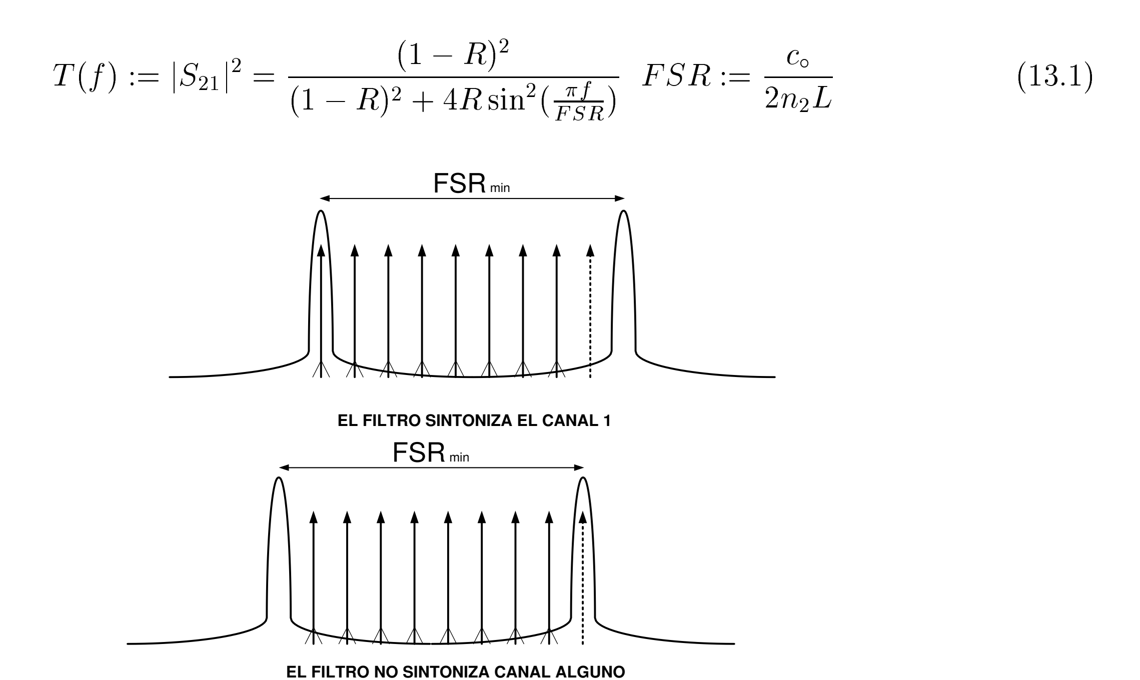


Figura 13.2: Rango espectral libre (FSR) mínimo del filtro para que el multiplexor "add & drop" funcione correctamente.

Los filtros de cavidad Fabry–Pérot usados en sistemas de multiplexación en longitud de onda (WDM), no se construyen con materiales semiconductores como los vistos en dicho capítulo ya que los coeficientes de reflexión en potencia de las superficies de entrada y salida deben ser considerablemente más altos (recordemos que en las cavidades semiconductoras los coeficientes de reflexión eran del orden de $R = 0{,}3$), siendo $R = 0{,}9$ un valor típico.

Estos filtros presentan coeficientes de transmisión en potencia máximos (para múltiplos enteros de la frecuencia FSR, denominada frecuencia espectral libre) y mínimos (para

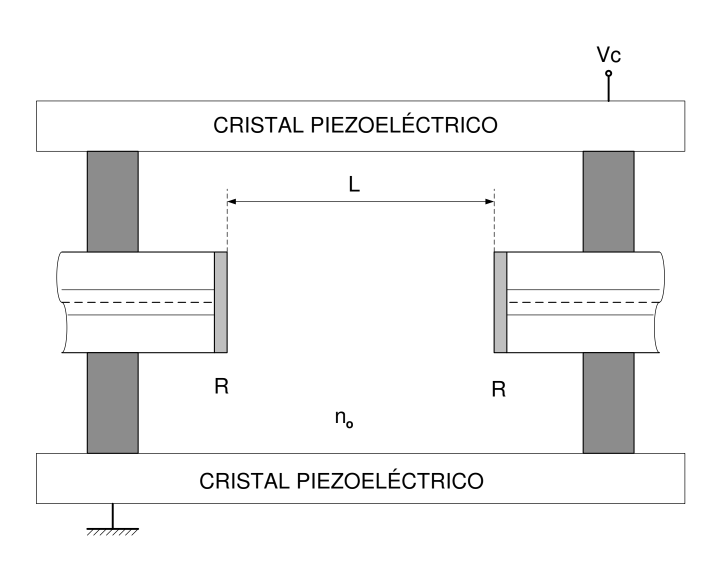


Figura 13.3: Implementación en tecnología todo fibra de un filtro de cavidad Fabry–Pérot sintonizable.

múltiplos semienteros de FSR) alternados. El valor máximo del coeficiente de transmisión es la unidad, mientras que el mínimo es $T_{min} = (\frac{1-R}{1+R})^2$. Normalmente hay que tratar las superficies de entrada y salida del filtro (por ejemplo, con la deposición de un material metálico) para que presenten un coeficiente de reflexión en potencia R elevado y que el filtro sea lo suficientemente selectivo. Una selectividad alta está directamente relacionada con un coeficiente de transmisión bajo a las frecuencias antirresonantes y, por extensión, en toda la banda de rechazo. Es fácil demostrar que la anchura espectral del filtro viene dada por:

$$FWHM := \frac{c_\circ}{\pi n_2 L} \sin^{-1}(\frac{1-R}{2\sqrt{R}}) \tag{13.2}$$

Los parámetros del filtro de cavidad Fabry–Pérot, R, FSR y $FWHM$, deben fijarse de la siguiente manera:

1. Debemos tomar un coeficiente de reflexión (R) que garantice que la atenuación en la banda de rechazo del filtro sea importante.

2. La anchura espectral del filtro en la banda de paso ($FWHM$) se elige de forma que la atenuación sea baja en todo el ancho de banda de la información transmitida (y, por tanto, no la distorsione), tal y como se muestra en la fig. {13.2}.

3. El rango espectral libre del filtro (FSR), debe ser superior a un valor mínimo que garantize que, en las condiciones más desfavorables, no seleccionamos un canal indeseado. En la fig. {13.2} hemos considerado que el filtro tiene la posibilidad de no sintonizar ningún canal (por lo que la banda de paso del filtro se centra a la izquierda del primer canal), y el siguiente lóbulo cae a la derecha del último canal (la flecha en trazo discontinuo indica que esa portadora no existe).

En la fig. {13.3} se muestra una posible realización física de un filtro de cavidad Fabry–Pérot sintonizable [72]. La longitud de la cavidad se puede modificar aplicando una

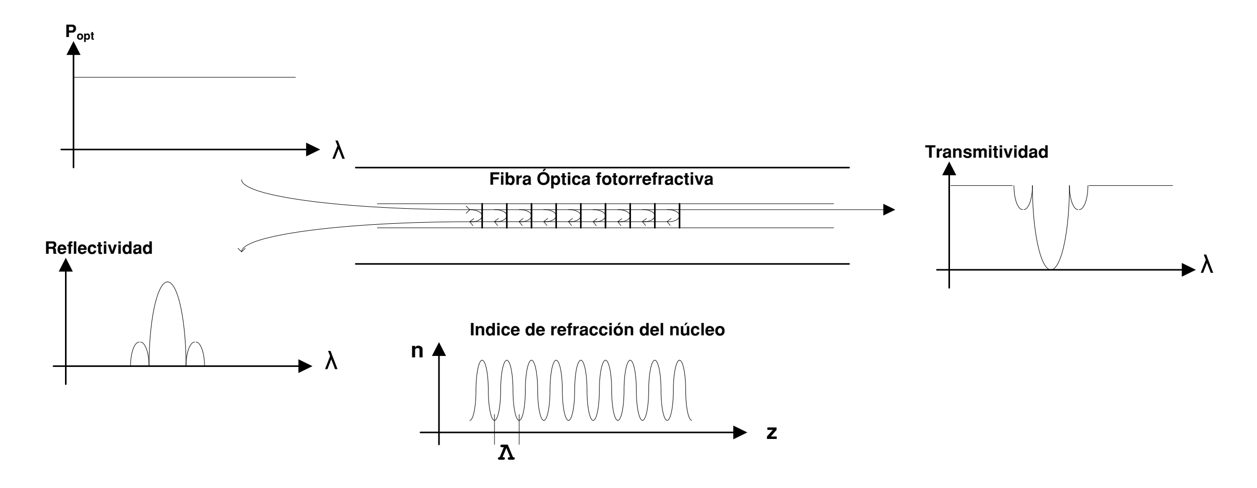


Figura 13.4: Fibras de difracción de Bragg en reflexión.

tensión al cristal piezoeléctrico. El cristal, al deformarse, modifica la distancia que separa las superficies reflectantes de las dos fibras monomodo.

13.3. Redes de difracción de Bragg en fibras (FBGs)

Las FBGs son fibras ópticas monomodo en las que, longitudinalmente, se introduce una perturbación senoidal del índice de refracción. Esta perturbación se "escribe" en la fibra, exponiendo al núcleo a una radiación ultravioleta que varía sinusoidalmente en intensidad a lo largo del eje de la fibra. En las zonas donde la intensidad luminosa de la radiación ultravioleta es nula, el índice de refracción del núcleo no sufre cambio alguno [2], mientras que en las iluminadas el índice de refracción local aumenta de manera proporcional a la potencia óptica, así como al tiempo de exposición. Este fenómeno recibe el nombre de fotosensibilidad y fue observado por primera vez en 1978 [39].

Los procesos físicos responsables de la fotosensibilidad del núcleo no están del todo establecidos, pero el modelo que cuenta con más aceptación es el de los centros de color. En este modelo, la radiación ultravioleta desencadena una serie de *reacciones fotoquímicas* que modifican las concentraciones de una serie de defectos ópticamente activos [7]. Por defecto, entendemos cualquier microestructura distinta a la de un tetraedro elemental *de sustitución*, esto es, en el que el silicio se ha sustituido por germanio [3] (la matriz regular del sílice parece no jugar un papel importante en la fotosensibilidad de la fibra). Estos defectos parecen estar asociados a deficiencias de oxígeno en los enlaces covalentes del $Ge0_2$. Como los picos de absorción de reactivos y productos se producen a longitudes de onda diferentes, se produce una variación del espectro de absorción que provoca, en consecuencia, una variación del índice de refracción ya que ambos parámetros están ligadas por medio de las relaciones integrales de Kramers–Kronig.

En la inmensa mayoría de las ocasiones, las FBGs se diseñan para funcionar en refle-

[2]Esto no es del todo cierto. Tras exponer la fibra a la radiación ultravioleta, se produce un incremento del índice medio (en todos los puntos el mismo) y otro local (proporcional a la intensidad luminosa).

[3]Recordemos que el núcleo de la fibra tiene un índice de refracción mayor que el de la corteza precisamente porque parte de losátomos de silicio se sustituyen por átomos de germanio.

xión, tal y como se muestra en la fig. {13.4}. Esto es así porque las funciones de transferencia en transmisión son, necesariamente, de fase mínima [61] (es decir, las características de módulo y fase de la función de transferencia tienen una relación biunívoca), lo que impone severas restricciones a la hora del diseño de los filtros [94].

El análisis de las FBGs (es decir, dado un perfil de variación del índice, deducir la función de transferencia correspondiente) es sencillo [60]. No hay más que aplicar el formalismo matricial visto en el capítulo 8, sustituyendo el perfil senoidal por uno rectangular cuyo primer armónico tenga la misma amplitud que el perfil senoidal. Por supuesto, también puede utilizarse el formalismo de modos acoplados. En ambos casos hay que utilizar métodos numéricos, por lo que en la subsección siguiente nos vamos a limitar a exponer los resultados de las simulaciones, aunque aconsejamos al lector que los compruebe por sí mismo realizando sus propias simulaciones.

El problema de la síntesis de las FBGs (dada una función de transferencia deseada, determinar el perfil de variación del índice) es algo más complejo, aunque también está resuelto. No se van a explicar los algoritmos de síntesis y referimos al lector interesado a la bibliografía adjunta [31]–[79]–[95].

Como ya hemos dicho, las aplicaciones fundamentales de las FBGs en comunicaciones ópticas (en sistemas de multiplexación en longitud de onda) son dos: la realización física de filtros paso banda en reflexión y la implementación de compensadores pasivos de la dispersión.

13.3.1. FBGs para la implmentación de filtros

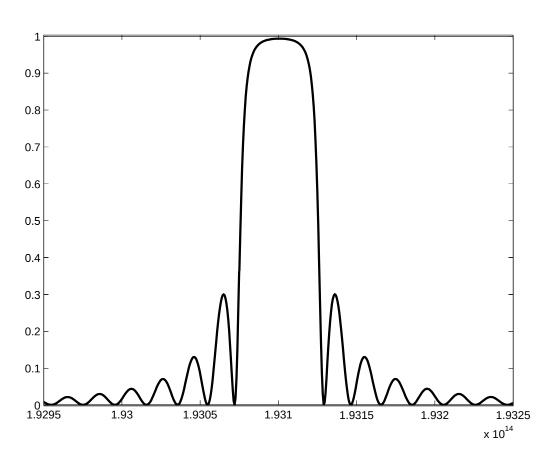
En la fig. {13.5} se muestran las características de módulo y de retardo de grupo [4] de una red de periodo uniforme [29]. En la fig. {13.6} se han representado los coeficientes de reflexión en potencia y retardo de grupos de una red de perfil apodizado, de perfil medio no constante [29]. Por último, en la fig. {13.7} se muestran las características de módulo y retardo de fase de una red apodizada con índice medio constante [29]. Vemos en la fig. {13.5} que hay lóbulos secundarios a ambos lados de la frecuencia central del filtro. Estos lóbulos aparecen por la variación brusca que sufre el índice a la entrada y la salida de la FBG, lo que provoca reflexiones internas por efecto Fabry–Pérot. En la fig. {13.6} sólo se presentan lóbulos secundarios a la derecha de la frecuencia central y en la última (fig. {13.7}) éstos desaparecen completamente.

En la fig. {13.8} se presentan, con los peridos magnificados, la variaciónes de índice asociadas a los tres filtros anteriormente citados.

13.3.2. Compensadores pasivos de la dispersión

Otra aplicación importante de las redes de disfracción de Bragg es la fabricación de dispositivos pasivos compensadores de la dispersión cromática. La dispersión cromática

[4] El retardo de grupo es la derivada de la característica de fase y representa el tiempo que transcurre entre la incidencia y la reflexión de un paquete de ondas.



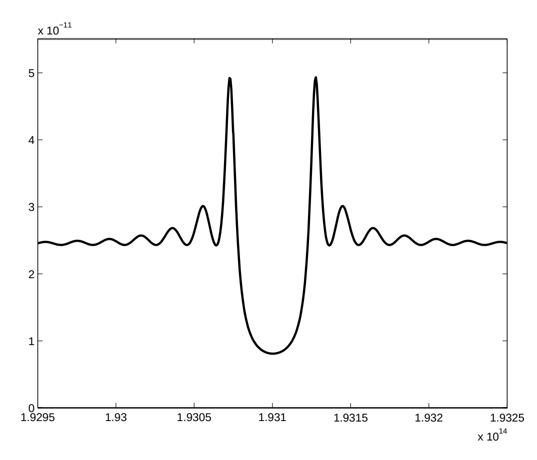
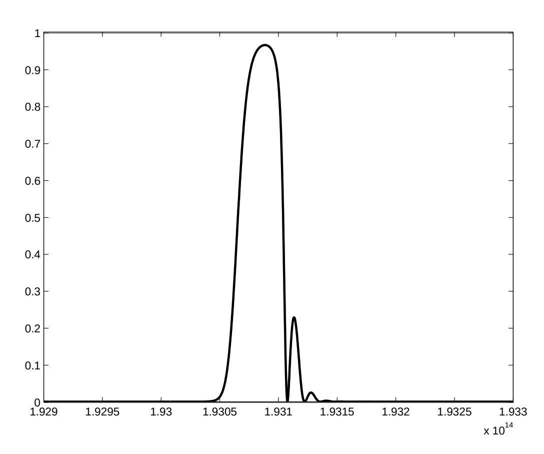


Figura 13.5: Red de difracción de Bragg uniforme. El índice de refracción medio es 1,452, el periodo de Bragg $\Lambda = 535nm$, la amplitud de la perturbación senoidal es 0.001 y la longitud de la fibra es $4mm$. Simulación realizada por Alejandro Carballar Rincón.



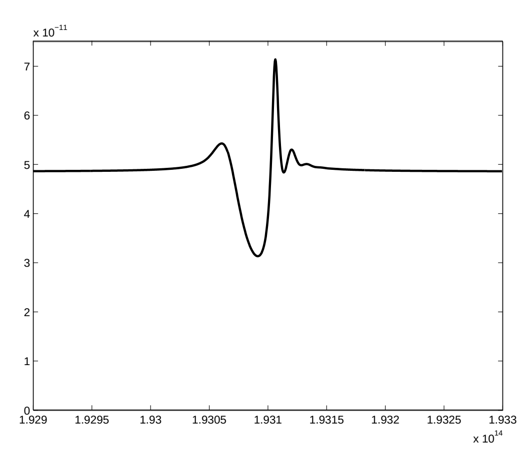
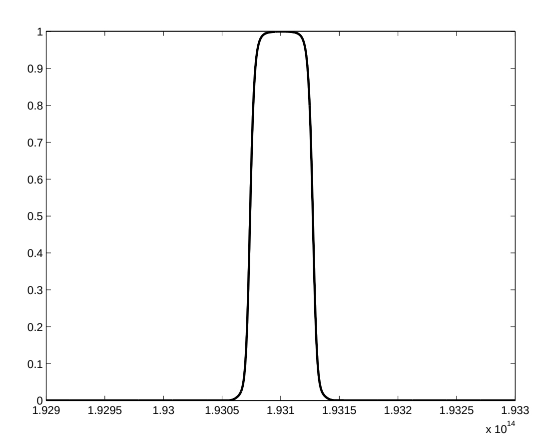


Figura 13.6: Red de difracción de Bragg apodizada con perfil de coseno alzado. El índice de refracción medio es 1,452, el periodo de Bragg $\Lambda = 534{,}47nm$, la amplitud de la perturbación senoidal tiene la expresión $0{,}0005\cos^2(\frac{\pi}{2} + \frac{\pi z}{L_{per}})$ y la longitud de la fibra es $L_{per} = 10mm$. Simulación realizada por Alejandro Carballar Rincón.



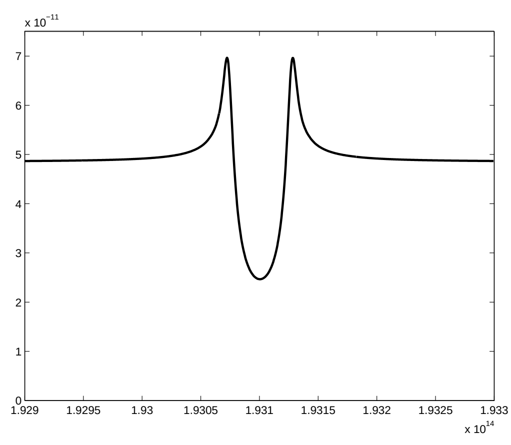


Figura 13.7: Red de difracción de Bragg apodizada con perfil de coseno de media cero. El índice de refracción medio es 1,452, el periodo de Bragg $\Lambda = 534{,}47nm$, la amplitud de la perturbación senoidal tiene la expresión $0{,}0004 + 0{,}0004\cos^2(\frac{\pi}{2} + \frac{\pi z}{L_{per}})$ y la longitud de la fibra es $L_{per} = 10mm$. Simulación realizada por Alejandro Carballar Rincón.

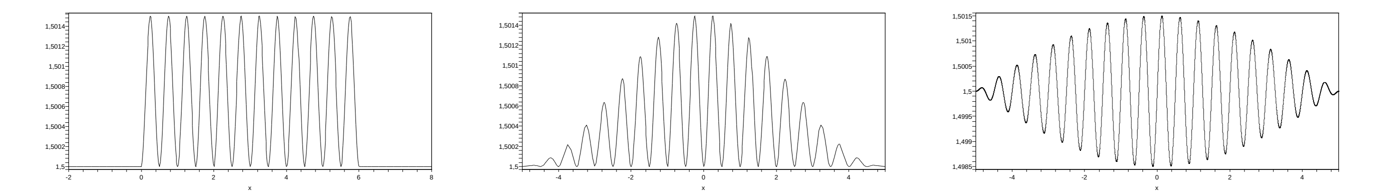


Figura 13.8: Perfiles (no están a escala) de las redes de Bragg uniforme (izquierda), apodizada (centro) y apadizada on valor medio constante (derecha).

de la fibra por la que se propagan las señales ópticas, hace que deteminadas componentes espectrales (las que "ven" un índice de grupo más bajo) se adelanten, ya que se propagan a una velocidad ligeramente mayor. Este efecto se puede compensar mediante una fibra de difracción de Bragg en la que el periodo de la perturbación varía linealmente a lo largo de la red, de manera que aquellas componentes que llegan más tarde (que se han retrasado) se reflejen al principio del dispositivo, de forma que el compensador de la dispersión introduzca, para estas componentes, una retardo de grupo nulo. Por contra, las componentes que más se han adelantado (las que han llegado antes), deben recorrer todo el dispositivo antes de encontrar la zona donde se cumple la condición de Bragg y reflejarse, de manera que se introduce una retardo de grupo que compensa *exactamente* el adelanto provocado por la dispersión cromática. Esto puede entenderse fácilmente viendo la fig. {13.9}.

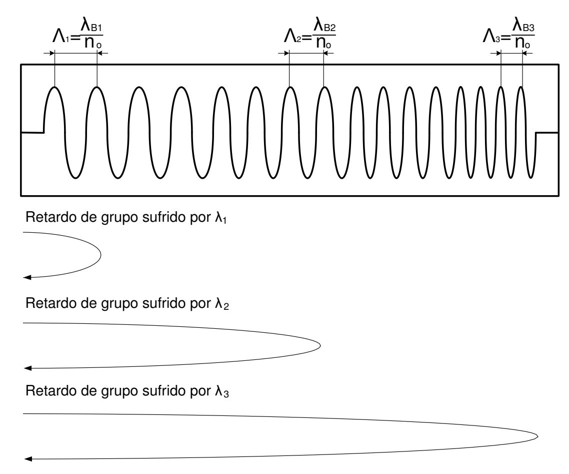


Figura 13.9: Red de difracción de Bragg con variación linel del periodo. Introduce un retardo de grupo que es proporcional (o inversamente proporcional si se coloca al revés) a la longitud de onda que llega al dispositivo. Este dispositivo es un compensador pasivo de la dispersión cromática.

Puesto que las fibras de difracción de Bragg funcionan en reflexión, el montaje del compensador de la dispersión incluye, necesariamente, un circulador óptico, tal y como se observa en la fig. {13.10}. Evidentemente, el periodo mínimo de Bragg se corresponde con la componente espectral de menor longitud de onda que llega al compensador (que estará al principio o al final del dispositivo dependiendo de que el signo de la dispersión cromática sea negativo o positivo), mientras que el periodo máximo se corresponde con la de mayor longitud de onda. La longitud del compensador es tal que el tiempo que transcurre en la ida y vuelta de la componente espectral que llega adelantada sea igual al retardo que sufre la componente más lenta. Por supuesto estamos suponiendo que el coeficiente de dispersión

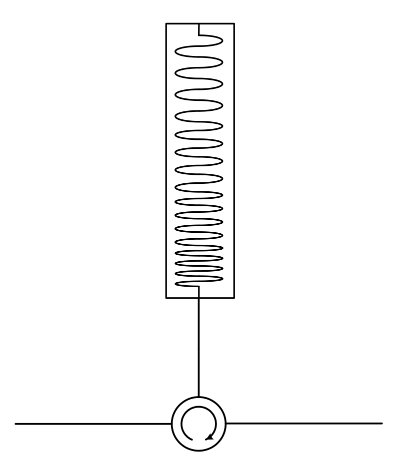


Figura 13.10: Compensador pasivo de la dispersión. Consta de un circulador y una rede de difracción de Bragg con variación lineal del periodo.

cromática permanece constante en todo el ancho de banda de las señales ópticas transmitidas, con lo que una variación lineal del periodo de Bragg es capaz de compensar este valor de dispersión cromática constante.

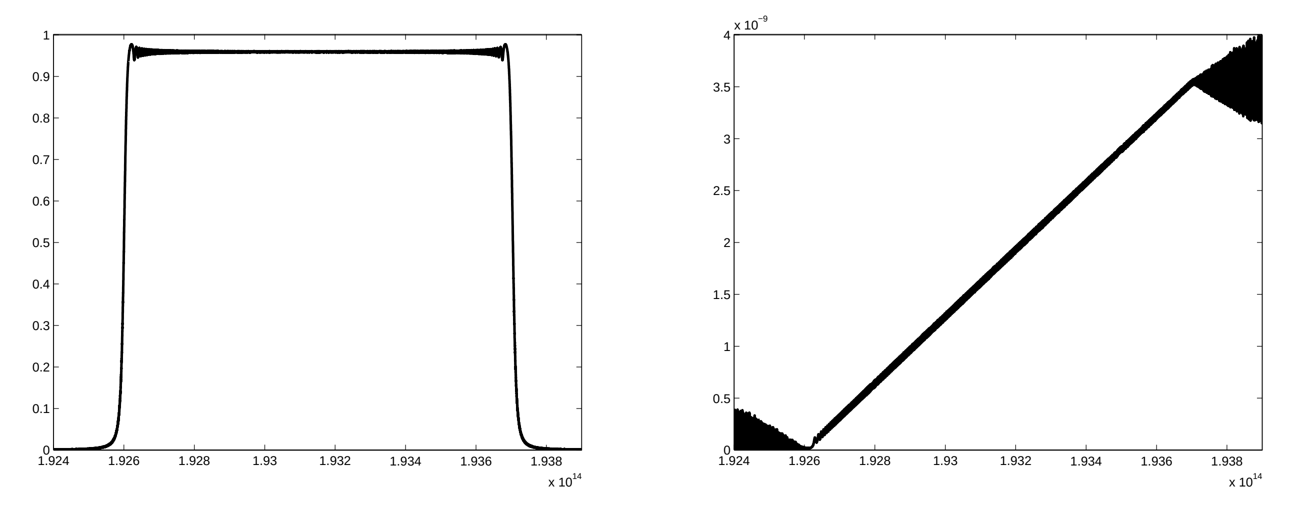


Figura 13.11: Compensador pasivo de la dispersión. Módulo a la izquierda y retardo de grupo a la derecha. Simulación realizada por Alejandro Carballar Rincón.

A continuación vamos a determinar la función de transferencia (coeficiente de reflexión y retardo de grupo) de un compensador pasivo de la dispersión. El ancho de banda de la señal óptica va desde los $192,6THz$ hasta los $193,7THz$. La longitud de la fibra es de $100Km$ y su coeficiente de dispersión cromática es $D = \frac{4ps}{nmKm}$. El índice de refracción medio de la red de difracción es $1,452$, mientras que la amplitud de la perturbación del índice es $\Delta n = 0{,}0004$. Realizando los cálculos correspondientes, resulta que $\Lambda_{min} = 534,83nm$, $\Lambda_{max} = 535,95nm$ y la longitud del compensador es $L = 0,365m$. Las simulaciones las vamos a hacer basándonos en el método matricial descrito en el capítulo 9, considerando perfiles rectangulares cuyo primer armónico coincide con la amplitud de la perturbación senoidal. Los resultados se muestran en las fig. {13.11}.

Problema 18: Se tiene un sistema WDM donde se transmite a las frecuencias f_1=192,6THz, f_2=192,7THz, f_3=192,8THz y f_4=192,9THz. Se quiere diseñar un filtro de cavidad Fabry–Pérot para que selecciones la frecuencia f_1. Variando la longitud de la cavidad, el filtro debe ser capaz de sintonizar cualquier otro canal, incluyendo el canal nulo (reflejar todas las frecuencias insidentes). En la banda de rechazo, se necesita un T_{min}=0,02. El índice de refracción del medio es n=1,5.

Solución: Debemos tomar una $FSR_{min} = 500GHz$ para tener la posibilidad de reflejar todos los canales, (el periodo de repetición de las bandas de paso del filtro es, precisamente, la FSR). Puesto que en la banda de rechazo necesitamos una $T_{min} = 0{,}02$ resulta:

$$0,02 = \left(\frac{1-R}{1+R}\right)^2 \Rightarrow R = 0,75$$

Para sintonizar el canal $f_1 = 196,6THz$ el cociente $\frac{f_1}{FSR}$ debe ser entero. Comprobémoslo:

$$k = \frac{f_1}{FSR} = 385,2$$

Este número no es entero, pero la FSR puede ser mayor que 500GHz, por los que tomamos $k' = 385$ y calculamos la FSR real:

$$\frac{c_\circ}{2nL} = FSR_{real} = \frac{f_1}{385} = 500,26GHz \Rightarrow L = 199,89\mu m$$

Ya podemos calcular la anchura del filtro:

$$FWHM = \frac{c_\circ}{\pi nL}\sin^{-1}\left(\frac{1-R}{2\sqrt{R}}\right) = 46,15GHz$$

Capítulo 14

Acopladores ópticos

Resumen: En este capítulo vamos a describir los acopladores ópticos.
Los acopladores son dispositivos pasivos que distribuyen una o varias señales ópticas de entrada entre varias salidas. Se fabrican tanto en tecnología integrada como en tecnología todo fibra y tienen un campo de aplicación muy importante en las redes ópticas pasivas (redes PON).

14.1. Principio de funcionamiento

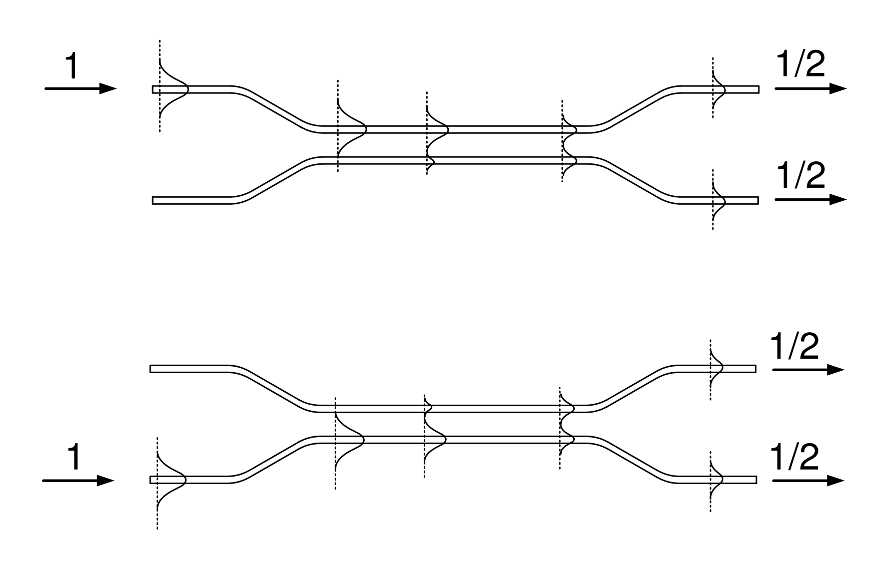


Figura 14.1: Acoplador direccional.

Los acopladores ópticos más utilizados, los denominados acopladores direccionales, se fabrican generalmente en tecnología de óptica integrada [65]–[75]. Normalmente se emplea el niobato de litio (NbO_3Li) como material sustrato sobre el que se realiza un dopado (con titanio, por ejemplo) en aquellas regiones cuyo índice de refracción se quiere aumentar, creándose, de esta forma, una estructura guíaonda de geometría planar. En la fig. {14.1} se muestran dos guiaondas (las zonas rectangulares representan las zonas de mayor índice) que se acercan mucho entre sí, de forma que el campo evanescente (la zona donde el campo electromagnético decrece exponencialmente) de la onda introducida por uno de los puertos, solapa parcialmente con el núcleo del otro guiaonda. Si estuviesen aislados el uno del otro, Los dos guiaondas soportarían, a la longitud de onda de trabajo, un único modo de propagación. La estructura mostrada se puede estudiar considerando que, al acercarse los dos núcleos, se produce una pequeña perturbación que acopla entre sí los fenómenos de propagación que, en ausencia de dicha perturbación, serían independientes el uno del otro. Puesto que hemos decidido, por su relativa complejidad, no introducir el formalismo de modos acoplados [103]–[65], no vamos a deducir las ecuaciones resultantes. Si lo hiciésemos y las integrásemos, comprobaríamos que, para una determinada longitud de la zona de acoplo L, se produciría una transferencia total de energía de un guiaonda al otro. Si incrementásemos la zona de acoplo en otra longitud L, se produciría el fenómeno inverso a éste y el intercambio se repetiría de manera periódica si la longitud de la zona de acoplo fuese infinita. En la fig. {14.1}, la longitud de acoplo es la necesaria para que se transfiera sólo la mitad de la potencia óptica.

Si hacemos una serie de consideraciones generales acerca del funcionamiento del dispositivo, podremos deducir que su matriz de scattering depende de un único parámetro [72]. En la fig. {14.2} hemos representado el acoplador direccional y hemos enumerado los cuatro puertos que lo constituyen. También hemos representado una matriz de scattering genérica

$$S_{ij} = \begin{bmatrix} 0 & 0 & S_{13} & S_{14} \\ 0 & 0 & S_{23} & S_{24} \\ S_{31} & S_{32} & 0 & 0 \\ S_{41} & S_{42} & 0 & 0 \end{bmatrix}$$

Figura 14.2: Matriz de scattering del acoplador direccional.

que describe su funcionamiento. Lo único que hemos tenido en cuenta al escribirla es que, en ausencia de reflexiones en los puertos, a partir de una onda que se propaga en un sentido no se puede generar otra que se propague en sentido contrario, de ahí los ceros que aparecen en la matriz. Si tenemos en cuenta que el dispositivo es recíproco, deben cumplirse las siguientes relaciones:

$$\begin{aligned} S_{13} &= S_{24} \\ S_{31} &= S_{42} \end{aligned} \tag{14.1}$$

Si despreciamos las pérdidas que se producen en el acoplador, tenemos que la matriz de scattering del acoplador direccional es unitaria, por lo que se verifican las siguientes igualdades:

$$\begin{aligned} S_{13}S_{13}^* + S_{14}S_{14}^* &= 1 \\ S_{23}S_{23}^* + S_{24}S_{24}^* &= 1 \\ S_{13}S_{23}^* + S_{14}S_{24}^* &= 0 \\ S_{31}S_{31}^* + S_{41}S_{41}^* &= 1 \\ S_{41}S_{41}^* + S_{42}S_{42}^* &= 1 \\ S_{31}S_{41}^* + S_{32}S_{42}^* &= 0 \end{aligned} \tag{14.2}$$

De donde deducimos que la matriz de scattering se puede escribir como:

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \sqrt{1-\alpha} & j\sqrt{\alpha} \\ 0 & 0 & j\sqrt{\alpha} & \sqrt{1-\alpha} \\ \sqrt{1-\alpha} & j\sqrt{\alpha} & 0 & 0 \\ j\sqrt{\alpha} & \sqrt{1-\alpha} & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \tag{14.3}$$

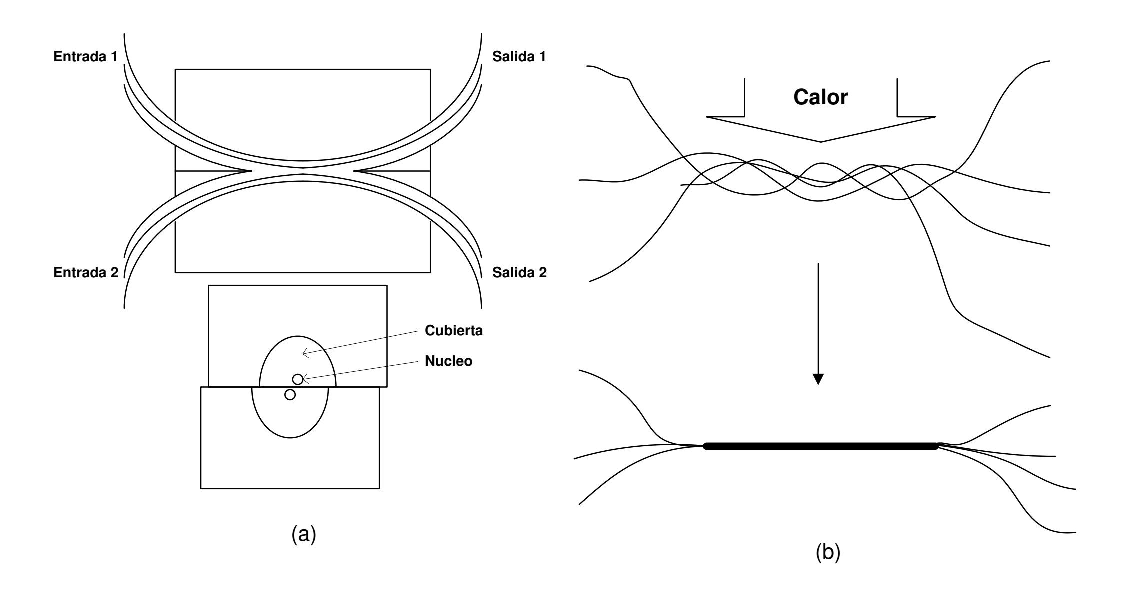


Figura 14.3: Acopladores ópticos fabricados en tecnología de fibra. De campo evanescente (izquierda) y fibras fusionadas (derecha)

El parámetro real α indica cómo se reparte la potencia entre cada uno de los dos puertos de salida y típicamente toma el valor de 0,5. Para calcular la longitud de la región de acoplo necesaria para obtener un valor de α determinado es necesario, evidentemente, resolver la ecuación de modos acoplados.

Los acopladores ópticos se fabrican también en tecnología de fibra [72], tal y como muestra la fig. {14.3}. El representado a la izquierda (cada una de las fibras soporta un único modo) tiene un coeficiente de acoplo variable ya que desplazando lateralmente una fibra con respecto a la otra, el solapamiento entre los dos modos guiados cambia. En el acoplador representado a la derecha, los núcleos de las distintas fibras están fundidos por lo que no se puede variar el coeficiente de acoplo.

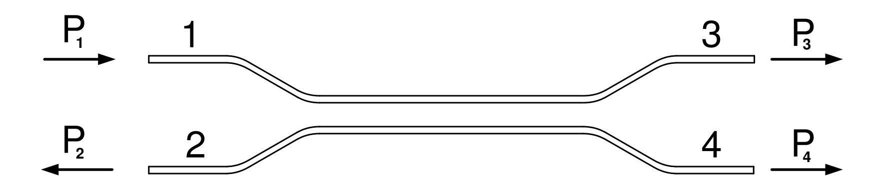


Figura 14.4: Potencias que salen por los distintos puertos del acoplador direccional tras excitarlo por la entrada "1".

14.2. Parámetros característicos

Hasta ahora hemos supuesto que el dispositivo se comporta de manera ideal, es decir, que no introduce pérdidas y que no hay propagación de ondas contradireccionales (se entiende que con los puertos adaptados). En la práctica esto no es así por lo que es necesario

definir una serie de parámetros que describan el comportamiento real de un acoplador óptico. Tomando como referencia la fig. {14.4}, definimos [72]:

1. **Constante de acoplo**: $k := \frac{P_4}{P_3+P_4}$ que, típicamente toma el valor 0,5.
2. **Pérdidas en exceso**: $\Gamma(dB) := 10\log_{10}(\frac{P_1}{P_3+P_4})$ que, típicamente vale 0,5 dB.
3. **Directividad**: $D(dB) := 10\log_{10}(\frac{P_2}{P_1})$ que, típicamente toma el valor de -50 dB.
4. **Pérdidas de inserción**: $M_i(dB) := 10\log_{10}(\frac{P_1}{P_i})$ i=3, 4 que, típicamente vale 3,8 dB.

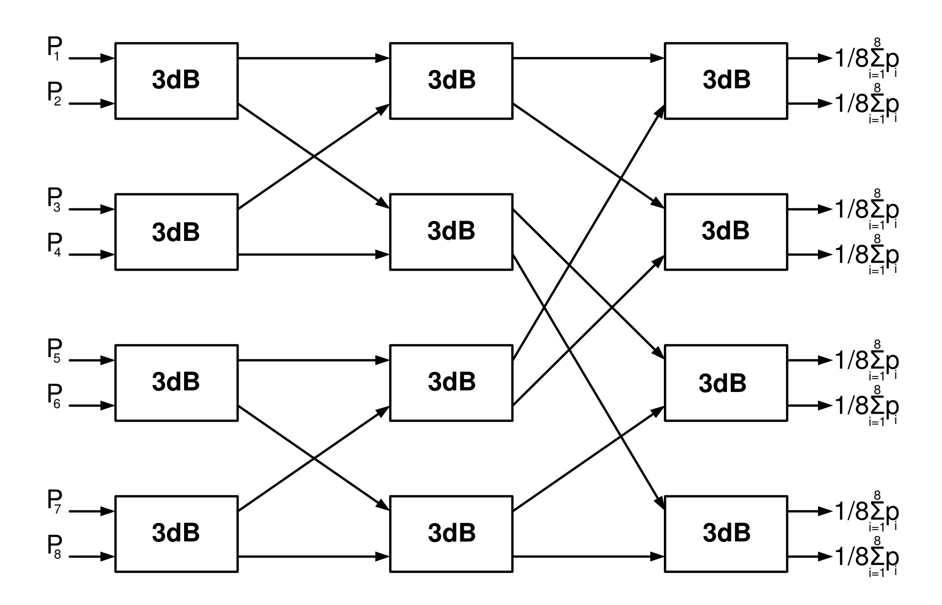


Figura 14.5: Acoplador en estrella 8 × 8.

14.3. Acopladores en estrella

Por acoplador en estrella entendemos un dispositivo óptico pasivo que, a partir de un conjunto de señales de entrada, las combina y distribuye entre varias salidas [72]. Normalmente se construyen a partir de acopladores direccionales de constante de acoplo 0,5. En la fig. {14.5} se muestra una acoplador en estrella 8×8 fabricado a partir de acopladores direccionales 2×2. Estos acopladores se utilizan mucho en redes ópticas pasivas [84], conocidas como redes PON (del inglés "passive optical networks").

Capítulo 15

Multiplexores y conmutadores ópticos

Resumen: En este capítulo vamos a describir otros dispositivos empleados en las redes DWDM: los multiplexores (y demultiplexores) y los conmutadores ópticos. Los primeros introducen (o separan) las distintas longitudes de onda que viajan simultáneamente por la fibra. Los segundos definen el encaminamiento de la información por la red óptica.

15.1. Multiplexores y demultiplexores

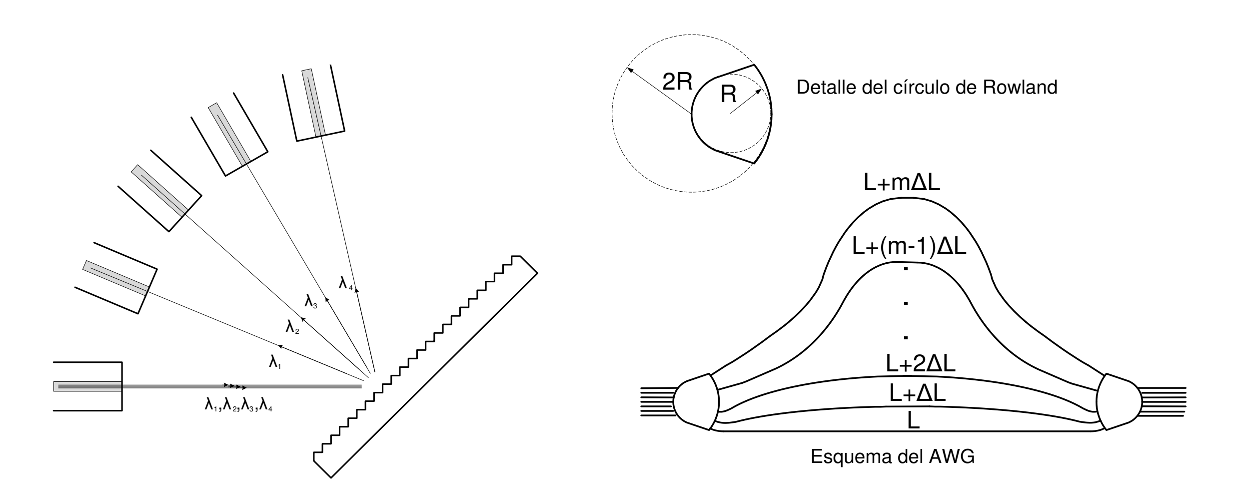


Figura 15.1: Multiplexor/demultiplexor fabricado en tecnología microóptica (izquierda). Multiplexor para encaminamiento estático (derecha).

Los multiplexores sirven para introducir varios canales, cada uno de ellos codificado a una longitud de onda distinta, en la misma fibra óptica, mientras que los demultiplexores realizan la función inversa. Este tipo de dispositivos se emplean, evidentemente, en sistemas DWDM y pueden fabricarse en distintas tecnologías. Así, por ejemplo, en la fig. {15.1} se representa, a la izquierda, un multiplexor/demultiplexor fabricado en tecnología microóptica [8]. Su principio de funcionamiento se basa en una red de difracción espacial en la que el ángulo de difracción varía de manera importante con respecto a la longitud de onda. Para ello es necesario que las perturbaciones en la superficie del espejo sean de un tamaño comparable (o más pequeñas) a las longitudes de onda de los canales a multiplexar o demultiplexar, ya que los ángulos de difracción son proporcionales a $\frac{\lambda_i}{d}$, donde d es la longitud de la perturbación existente en el espejo.

En los sistemas DWDM se emplean multiplexores más complejos para definir tablas de encaminamiento estático entre los nodos de la red. En la fig. {15.1} se muestra, a la derecha, un AWG (del inglés "Arrayed Waveguide Grating") o WGR (también del inglés "Waveguide Grating Router") [8]–[65]–[96]. Como se observa, el AWG está compuesto por un acoplador de entrada, m guiaondas, cada uno de los cuales introduce un retardo constante con respecto al anterior (ya que es ΔL metros más largo que el anterior) y un acoplador de salida. Los acopladores de entrada y salida son "círculos de Rowland". El círculo de Rowland de la izquierda realiza m copias de la información que de las n fibras entrantes. Los guiaondas introducen un retardo constante entre dos entradas consecutivas y, finalmente, esta información se vuelve a combinar en el acoplador de salida. Ahora, para cada canal y longitud de onda, las condiciones de interferencia constructiva dentro del círculo de Rowland de la derecha no tienen por qué darse (y de hecho en general no es así) para la misma fibra que a la entrada.

Una posible tabla de encaminamiento sería la mostada en la fig. {15.2} (parte derecha). El dispositivo se trata de un demultiplexor–multiplexor de cuatro entradas y cuatro salidas y su funcionamiento es el siguiente: saca la longitud de onda λ_1 por la misma fibra

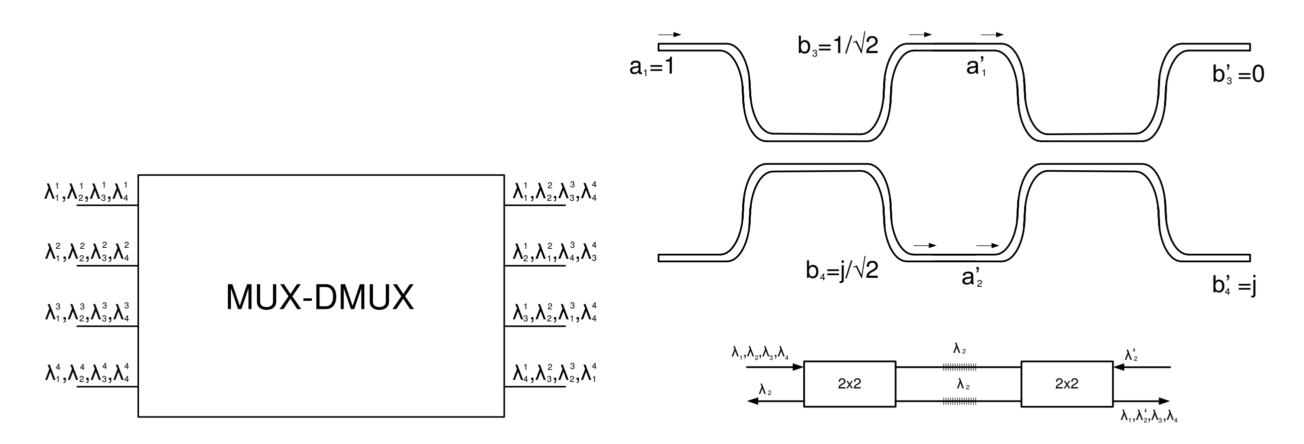


Figura 15.2: Ejemplo de tabla de encaminamiento de un AWG (izquierda). Multiplexor add and drop fabricado con acopladores direccionales (derecha).

por la que entra, la longitud de onda λ_2 por la siguiente siguiente, la información codificada en λ_3 por la fibra desplazada dos posiciones y, por último, la longitud de onda λ_4 por la que está desplazada tres posiciones.

Otro tipo de multiplexor, que ya hemos estudiado en un capítulo anterior, es el multiplexor "add and drop". En la parte derecha de la fig. {15.2} se muestra otra posible implementación de este multiplexor que, básicamente, consiste en dos acopladores de $k = 50\,\%$ implementados en óptica integrada [8]. Además, en los guiaondas que conectan los dos acopladores, se deposita un reflector de Bragg sintonizado a la longitud de onda que se quiere sustituir. Para entender su funcionamiento lo único que tenemos que hacer es calcular la matriz de reparto de dos acopladores colocados en cascada:

$$\begin{bmatrix} b_1' \\ b_2' \\ b_3' \\ b_4' \end{bmatrix} = (\frac{1}{\sqrt{2}})^2 \begin{bmatrix} 0 & 0 & \frac{\sqrt{2}}{2} & j\frac{\sqrt{2}}{2} \\ 0 & 0 & j\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & j\frac{\sqrt{2}}{2} & 0 & 0 \\ j\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & \frac{\sqrt{2}}{2} & j\frac{\sqrt{2}}{2} \\ 0 & 0 & j\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & j\frac{\sqrt{2}}{2} & 0 & 0 \\ j\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ j \end{bmatrix} \quad (15.1)$$

Vemos que lo que entra por el puerto "1" sale por el "4" desfasado 90°. Teniendo en cuenta este resultado, resulta fácil deducir que el funcionamiento de este multiplexor es el representado en la fig. {15.2}.

15.2. Conmutadores ópticos

Los conmutadores ópticos se utilizan en sistemas DWDM con distintas finalidades. En función de la aplicación, los tiempos de conmutación han de ser de un determinado orden de magnitud [8]–[84]:

- **Provisión de nuevos enlaces.** De esta manera podemos reconfigurar la red óptica "off-line", es decir,cuando no está en uso. En este caso los tiempos de conmutación están en el rango de 1–10 ms. Suelen fabricarse en tecnología microóptica y no son más que una matriz de microespejos y su conmutación consite en mover (girar) los espejos.

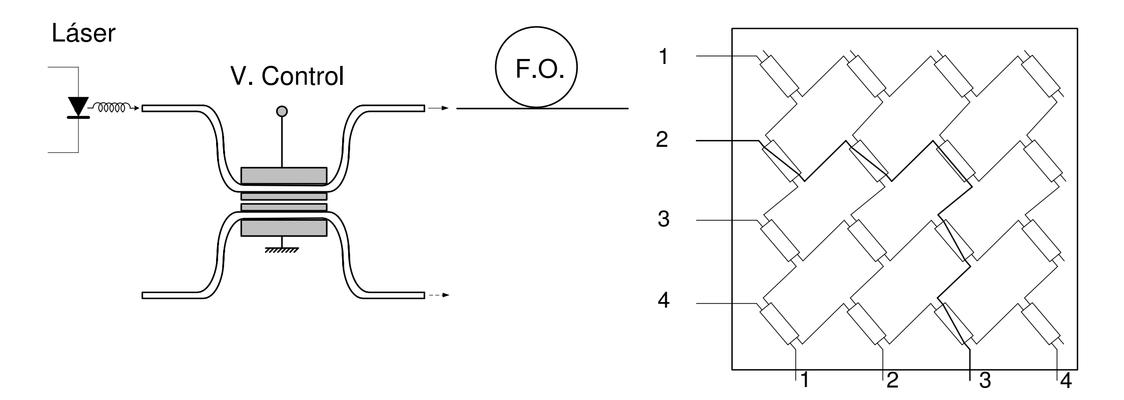


Figura 15.3: Modulador electroóptico en óptica integrada (izquierda). Conmutador rápido crossbar (derecha).

Son relativamente baratos, por lo que pueden fabricarse conmutadores con un número muy elevado de entradas y salidas.

- **Conmutación para protección.** Se realiza cuando una fibra falla (por rotura, por ejemplo) y en el cable óptico hay fibras no usadas (llamadas fibras oscuras) que son redundantes. Los tiempos de conmutación deben estar en el rango de 1–10μs.

- **Conumutación de paquetes ópticos.** Las redes DWDM tienden a incorporar multiplexación temporal en el dominio óptico (OTDM) con el objeto de implimentar en el dominio óptico las capas de la red correspondientes al encaminamiento de la información. Las redes existentes no incluyen esta posibilidad pero se cree que cuando se consigan fabricar de manera eficiente conversores en longitud de onda (WLCs del inglés "wavelength converters") esta técnica se llevará a la práctica. Los tiempos de conmutación estimados son del orden de 1ns.

- **Moduladores externos.** Los lásers que se fabrican en la actualidad no se pueden modular por inyección de corriente por encima de unos pocos Gb/s. Las tasas binarias de 10 y 40 Gb/s (estándares en sistemas DWDM, al menos la primera) se consigue empleando moduladores externos, es decir, dispositivos que se pueden hacer transparentes u opacos en función de una tensión aplicada. Existen dos tipos de moduladores externos: moduladores de electroabsorción y acopladores por efecto electroóptico. Los primeros son dispositivos a base de pozos cuánticos, cuyo estudio no se aborda en este libro ya que su principio de funcionamiento se basa en un efecto cuántico. Los segundos son acopladores por campo evanescente en los que se puede variar rápidamente la longitud efectiva de acoplo por efecto electroóptico. Los tiempos de conumutación son del orden de 10ps. En la fig. {15.3} se ha representado un modulador de este tipo.

Cuando se necesitan conmutadores de múltiples entradas y salidas con tiempos de conumutación muy cortos, no queda más remedio que que fabricar matrices de acopladores 2×2 que funcionen por efecto electroóptico. Así, por ejemplo, se muestra en la fig. {15.3} un conmutador de configuración "crossbar". Este dispositivo tiene la ventaja de que no se bloquea nunca y el inconveniente de requerir un númer N^2 de acopladores 2×2, donde N es el número de fibras de entrada y de salida.

Capítulo 16

Efectos no lineales

Resumen: En los sistemas DWDM se propagan centenares de canales multiplexados en longitud de onda, cada uno de los cuales transmite una tasa binaria del orden de algunas decenas de Gb/s. Las distancias a las que se transmite la información pueden ser del orden de algunos miles de kilómetros (por ejemplo en las comunicaciones transoceánicas). La conjunción de estos dos hechos: potencias ópticas muy altas y distancias recorridas muy elevadas, hacen que sea preciso tener en cuenta las interacciones no lineales entre los pulsos luminosos. Como regla general, los efectos no lineales son procesos no deseados que degradan la calidad del enlace. En la bibliografía se dan criterios de diseño que limitan las potencias máximas que se pueden transmitir por cada canal en función de la no linealidad considerada. En este capítulo nos vamos a limitar a enumerar y explicar el origen físico de estos procesos.

16.1. Introducción

La óptica no lineal estudia los fenómenos en los que las propiedades del medio material se ven modificadas por la presencia de una luz muy intensa. Típicamente sólo la luz láser es lo suficientemente potente como para producir efectos no lineales.

El sílice es un medio material poco propenso a sufrir este tipo de fenómenos [18]–[45] de manera que en los sistemas donde se transmite un único canal y, a menos que se hayan diseñado para que propaguen solitones, los efectos no lineales son poco importantes. La situación es, sin embargo, muy distinta en los sistemas DWDM, donde se propagan centenares de canales a lo largo de centenares o incluso miles de kilómetros. Es la conjunción de estos dos fenómenos (potencias muy elevadas propagándose a lo largo de enormes distancias) la que provoca la aparición de estos fenómenos que, en general, no son deseados y que degradan las prestaciones del enlace [26].

$$\ddot{x}+\gamma\dot{x}+\omega_o^2 x=-qE(t)/m$$

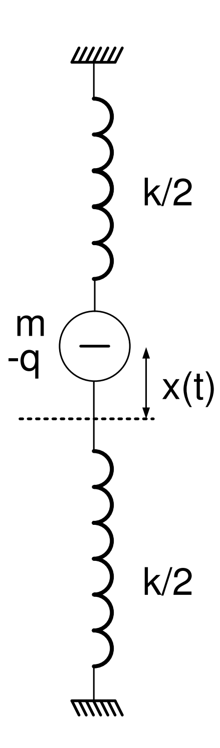


Figura 16.1: Modelo de Lorentz de un átomo.

16.2. Origen microscópico

En el modelo mecánico del átomo representado en la fig. {16.1} suponemos que la fuerza electromagnética que mantiene al electrón orbitando alrededor del núcleo es $F = -kx(t) = -m\omega_o^2 x(t)$. Esto es cierto siempre que el desplazamiento $x(t)$ del electrón con respecto a su posición de equilibrio sea pequeño. Si el campo eléctrico $E(t)$ se hace muy

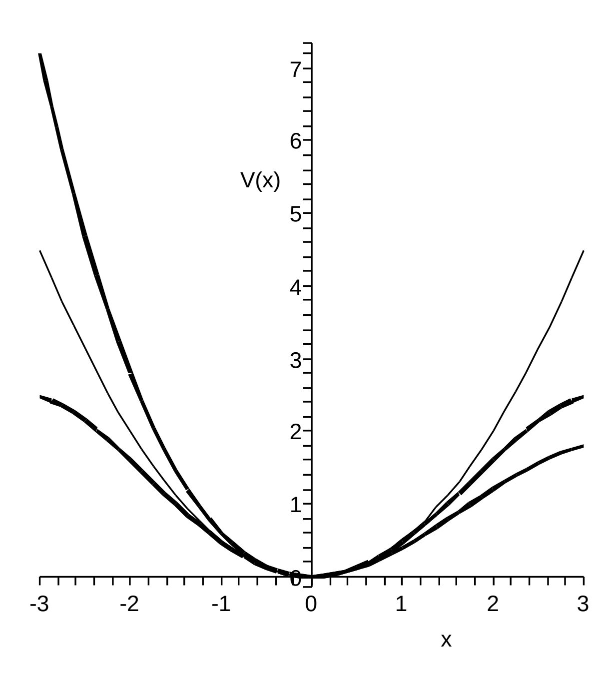


Figura 16.2: Desviaciones del potencial respecto al parabólico.

grande porque la luz es muy intensa, esta hipótesis deja de ser cierta, la fuerza que ejercen los muelles dejan de ser proporcionales al desplazamiento y ya no derivan, por tanto, del campo potencial parabólico $V(x) = \frac{1}{2}m\omega_\circ^2 x^2(t)$.

En la fig. {16.2} se han representado en trazo grueso dos tipos de desviaciones frente al potencial parabólico: la curva asimétrica representa una no linealidad propia de un medio material sin simetría de inversión (no es el caso del sílice) mientras que la asimétrica es propia de un medio **con** simetría de inversión (es el caso del sílice) [18]. Sólo vamos a considerar este último caso.

Para poder estudiar analíticamente lo que ocurre en presencia de campos electromagnéticos intensos debemos modificar la ecuación diferencial de Lorentz e introducir un término no lineal ad–hoc pequeño. La ecuación resultante es [18]:

$$\ddot{x} + \gamma\dot{x} + \omega_\circ^2 x - bx^3 = -\frac{qE(t)}{m} \tag{16.1}$$

Si resolvemos la ecuación diferencial (16.1) aplicando un método perturbativo, llegamos a la conclusión de que la polarización microscópica provocada por un campo eléctrico incidente senoidal tiene un aspecto parecido la mostrado en la fig. {16.3}. A la hora de calcular el momento dipolar macroscópico no podemos, evidentemente, pasar al dominio espectral ya que las ecuaciones no son lineales. La polarización macroscópica la podemos expresar como:

$$\vec{P}(t) = \int_{-\infty}^{t} \chi^{1)}(t-\tau)\vec{E}(\tau)d\tau + \int_{-\infty}^{t}\int_{-\infty}^{t}\int_{-\infty}^{t} \chi^{3)}(t-\tau_1; t-\tau_2; t-\tau_3)\vec{E}(\tau_1)\vec{E}(\tau_2)\vec{E}(\tau_3)d\tau_1 d\tau_2 d\tau_3$$

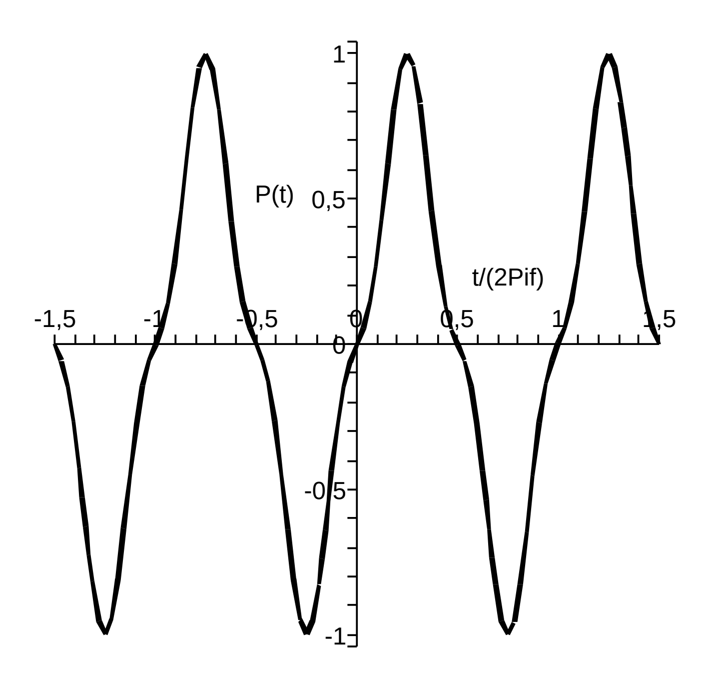


Figura 16.3: Momento dipolar microscópico no lineal (exagerado) provocado por la ecuación (16.1).

Ahora vamos a determinar el tensor de susceptibilidad de tercer orden $\chi^{3)}$ en una serie de situaciones que tienen especial importancia para nosotros, puesto que que describen fenómenos ópticos no lineales que se producen en los sistemas DWDM.

16.3. Efecto Kerr

Vamos a resolver la ecuación (16.1) utilizando un método perturbativo, ya que asumimos que $bx^2 \ll \omega_\circ^2$ [18]. Vamos a suponer que por la fibra se propaga *una única* portadora cuya pulsación es mucho menor que la pulsación característica del oscilador de Lorentz, es decir, $\omega \ll \omega_\circ$, por lo que el número de dipolos en resonancia es muy reducido y por tanto el tiempo medio entre colisiones muy alto ($T_\circ \simeq \infty$). Además, para simplificar las ecuaciones vamos a despreciar el término de fricción $\gamma\dot{x}$.

1. Aproximación de orden 1. No tenemos en cuenta el término no lineal bx^3 con lo que nos queda:

$$\ddot{x}_1 + \omega_\circ^2 x_1 = -\frac{qE(\omega)}{m}\cos(\omega t) \Rightarrow x_1(t) = -\frac{q/mE(\omega)}{\omega_\circ^2 - \omega^2}\cos(\omega t) \Rightarrow$$

$$\chi^{1)} = \frac{Nq^2}{dV\epsilon_\circ m(\omega_\circ^2 - \omega^2)} \tag{16.2}$$

2. Aproximación de orden 2. Consideramos la ecuación completa, sustituyendo en el término

de la derecha la aproximación de orden 1 y resolvemos en x_2:

$$\ddot{x}_2 + \omega_\circ^2 x_2 = -\frac{qE(\omega)}{m}\cos(\omega t) + bx_1^3 \tag{16.3}$$

$$\left.\begin{array}{r} \ddot{x}_2 + \omega_\circ^2 x_2 = -\dfrac{qE(\omega)}{m}\cos(\omega t) + bx_1^3 \\ x_1 = -\dfrac{q/mE(\omega)}{\omega_\circ^2 - \omega^2}\cos(\omega t) \end{array}\right\} \Rightarrow$$

$$x_2(t) = -\frac{qE(\omega)}{m(\omega_\circ^2 - \omega^2)}\cos(\omega t) - \frac{3q^3bE^3(\omega)}{4m^3(\omega_\circ^2 - \omega^2)^4}\cos(\omega t) - \frac{3q^3bE^3(\omega)}{4m^3(\omega_\circ^2 - \omega^2)^4}\cos(3\omega t) \tag{16.4}$$

Las oscilaciones asociadas al primer término son las responsables del momento dipolar lineal. El segundo término es el responsable de que aparezca un efecto no lineal que recibe el nombre de *efecto Kerr* y que consiste, tal y como probaremos más adelante, en un cambio en el índice de refracción del medio inducido por la propia potencia óptica que se propaga por la fibra. El tercer sumando describe la generación de un tercer armónico (una onda luminosa de pulsación triple) que no se produce realmente en fibras ópticas [1].

Teniendo en cienta que la pulsación es mucho menor que la pulsación característica del oscilador de Lorentz, es decir, $\omega \ll \omega_\circ$, resulta que la susceptibilidad dieléctrica y la polarización no lineal asociadas al efecto Kerr tienen la siguiente expresión:

$$\left.\begin{array}{r} \chi_K^{3)} \simeq \dfrac{3Nq^4b}{dV4m^3\omega_\circ^8} \\ P_K(\omega) = \chi_K^{3)}|E(\omega)|^2E(\omega) \end{array}\right\} \tag{16.5}$$

Lo que realmente nos interesa es calcular la variación producida en el valor de la susceptibilidad dieléctrica (o en el índice de refracción) por causa del efecto Kerr. Teniendo en cuenta que la relación que hay entre la intensidad luminosa y el campo eléctrico es $I = \frac{|E(\omega)|^2}{2\eta}$, donde η es la impedancia del medio, resulta [87]:

$$\left.\begin{array}{r} \Delta\chi = \dfrac{P_K(\omega)}{\epsilon_\circ} = \dfrac{2\eta\chi_K^{3)}}{\epsilon_\circ}I \\ n^2 = 1 + \chi \Rightarrow \Delta\chi = 2n\Delta n \end{array}\right\} \Rightarrow n(I) = n + n_2 I; \quad n_2 = \frac{\eta_\circ}{n^2\epsilon_\circ}\chi_K^{3)} \tag{16.6}$$

[1] El motivo es que no se verifica la condición de acuerdo de fase. En el apartado dedicado al mezclado de cuatro ondas, explicaremos en qué consiste esta condición.

16.4. Fenómenos no lineales asociados al efecto Kerr

16.4.1. Automodulación de fase (SPM)

Como consecuencia del efecto Kerr, un pulso óptico de intensidad luminosa alta que se propaga por una fibra, sufre un proceso de automudulación de fase (SPM del inglés "Self phase modulation"). Aquellas zonas del pulso donde la intensidad luminosa es mayor, "ven" un incremento local del índice de refracción cuyo valor es $\Delta n = n_2 I$, con lo que sufren un retardo de fase con respecto a las zonas donde la intensidad es menor. Si la envolvente del pulso tiene una potencia óptica P constante, la sección transversal de la fibra es A y la longitud de la fibra es L, entonces el retardo de fase es el mismo en todos los puntos del pulso y su valor sería $\Delta\varphi = 2\pi n_2 \frac{PL}{\lambda_\circ A}$. Evidentemente esta no es la situación habitual en comunicaciones ópticas, ya que la dispersión de la fibra suavizaría las transiciones de una envolvente rectangular.

En la fig. {16.4} se ha representado un pulso de envolvente Gaussiana, el desfase acumulado y el chirp inducido [47]. Vemos que la automodulación de fase produce una distribución no uniforme de las componentes espectrales a lo largo de la envolvente. En la figura no se ha simulado el efecto de la dispersión cromática de la fibra. Si la tenemos en cuenta, el chirp inducido hace que, en la mayor parte de las ocasiones, la dispersión cromática distorsione mucho la forma del pulso, por lo que, en general, se debe trabajar con potencias ópticas lo suficientemente débiles como para que no haya una automodulación de fase apreciable. La automodulación de fase es, pues, un fenómeno indeseado en los sistemas DWDM que debe evitarse. Sin embargo, si los sistemas se diseñan cuidadosamente, de manera que los pulsos tengan una envolvente muy especial (las llamadas ondas solitarias o, "solitones") la SPM combinada con una dispersión cromática negativa de un valor determinado hace que los pulsos se propaguen sin distorsión, es decir, sin sufrir dispersión alguna [46].

En la fig. {16.4} vemos que al principio del pulso se acumulan componentes espectrales de baja frecuencia, mientras que al final están las componentes de mayor frecuencia. En definitiva, la SPM induce un coeficiente de chirp ($C_\circ$) *negativo* en el pulso. Si la dispersión cromática de la fibra es negativa (esta es la situación habitual para la mayor parte los sistemas en tercera ventana) las componentes espectrales de menor frecuencia (que van al principio del pulso) se propagan con menor velocidad que las de mayor frecuencia (que van al final del pulso), por lo que se produce una "compresión" del pulso. En el capítulo 2 ya estudiamos un fenómeno parecido. La diferencia estriba en que ahora el fenómeno de compresión no se produce por la presencia de un chirp inicial en el pulso, sino por la automodulación de fase. Esta tendencia que tiene el pulso a comprimirse puede compensar completamente la dispersión cromática (que tiende a ensanchar el pulso) si la envolvente tiene una forma muy particular: la del "solitón" de primer orden [2]. En la actualidad existen algunos sistemas comerciales basados en una variación de este principio conocido como propagación de solitones con gestión de la dispersión [7]–[35]–[46].

[2] La ecuación no lineal de Schrödinger, que es la que describe la propagación de solitones por fibras ópticas, puede integrarse aplicanso el método del scattering inverso [100]–[108].

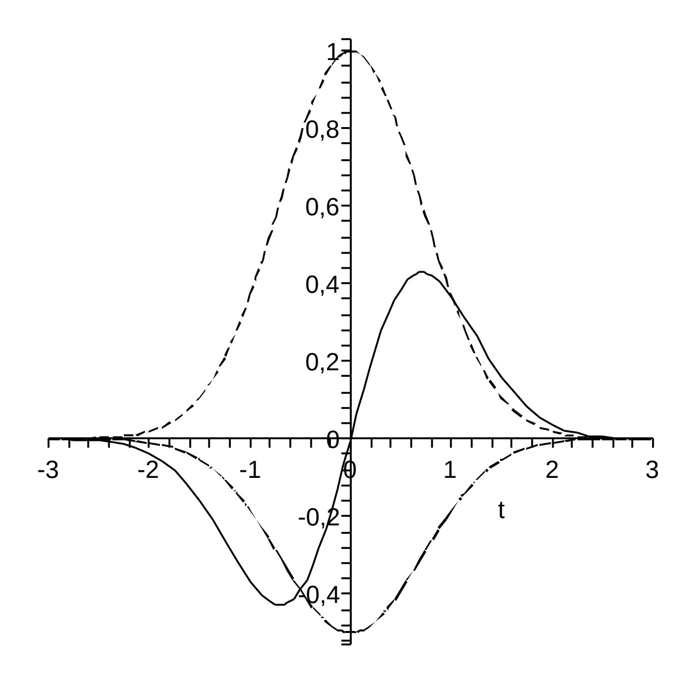


Figura 16.4: Chirp producido en un pulso Gaussiano por la automodulación de fase. La envolvente se ha representado en trazo discontinuo. El desfase es $\Delta\Phi = \frac{2\pi n_2 L}{\lambda_\circ A}P(t)$, mientras que el chirp (que se ha representado en trazo más fino) es $\Delta\omega = \frac{d\Delta\Phi}{dt}$.

16.4.2. Modulación de fase cruzada (XPM)

La modulación de fase cruzada (XPM del inglés "Cross phase modulation") aparece, también, como consecuencia del efecto Kerr. Ocurren en sistemas DWDM, en los que se produce una variación no lineal de la fase en los pulsos de un canal determinado como consecuencia de la potencia óptica que circula por el resto de los canales [6]. La expresión de la XPM es [26]:

$$\Delta\Phi_m(t) = \frac{2\pi n_2 L}{\lambda_\circ A}[P_{\circ m}(t) + 2\sum_{i\neq m} P_{\circ i}(t)] \tag{16.7}$$

El primer sumando representa la SPM, mientras que el segundo es la XPM. La situación peor se da cuando en el resto de los canales se propaga un pulso (es decir, un "1" lógico) que solapa totalmente con el pulso que se está propagando por el canal considerado. En esta situación la variación de fase producida por la modulación de fase cruzada vale:

$$\Delta\Phi_m(t) = \frac{2\pi n_2 L}{\lambda_\circ A}P_{\circ i}(t)(2M-1) \tag{16.8}$$

Evidentemente al evaluar esta expresión hemos sido excesivamente pesimistas, ya que hemos supuesto que todos los canales se propagan con igual velocidad. Esto es cierto sólo

si la dispersión cromática de la fibra es nula. Para reducir los problemas provocados por la XPM se utilizan fibras que tiene una dispersión cromática residual de unos 5ps/nmKm y se toma un espaciamiento entre canales lo suficientemente alto como para que la velocidad de propagación de los canales adyacentes sea considerablemente distinta [26]. De esta manera se consiguen paliar los problemas planteados por la modulación de fase cruzada, ya que los pulsos se cruzan con velocidades relativas bastante altas. Otra manera de minimizar la XPM es trabajar con fibras de dispersión cromática alta, pero alternando dispersiones positivas con negativas (iguales en valor absoluto), de manera que las velocidades relativas de propagación de los pulsos canales adyacentes son grandes.

16.5. Mezclado de cuatro ondas (FWM)

En el apartado anterior determinamos la susceptibilidad dieléctrica de tercer orden para el caso de una única onda incidente. Ahora vamos a considerar que se propagan tres ondas de pulsaciones ω_1, ω_2 y ω_3. En este caso nos queda (ahora vamos a representar las ondas por medio de fasores) [6]:

1. Aproximación de orden 1. No tenemos en cuenta el término no lineal bx^3 con lo que nos queda:

$$\ddot{x}_1 + \omega_\circ^2 x_1 = -\frac{qE(\omega_1)}{m}\frac{e^{j\omega_1 t}+e^{-j\omega_1 t}}{2} - \frac{qE(\omega_2)}{m}\frac{e^{j\omega_2 t}+e^{-j\omega_2 t}}{2} - \frac{qE(\omega_3)}{m}\frac{e^{j\omega_2 t}+e^{-j\omega_2 t}}{2} \Rightarrow$$

$$x_1(t) = -\frac{q/mE(\omega_1)}{\omega_\circ^2-\omega_1^2}\frac{e^{j\omega_1 t}+e^{-j\omega_1 t}}{2} - \frac{q/mE(\omega_2)}{\omega_\circ^2-\omega_2^2}\frac{e^{j\omega_2 t}+e^{-j\omega_2 t}}{2} - \frac{q/mE(\omega_3)}{\omega_\circ^2-\omega_3^2}\frac{e^{j\omega_3 t}+e^{-j\omega_3 t}}{2}$$

2. Aproximación de orden 2. Consideramos la ecuación completa, sustituyendo en el término de la derecha la aproximación de orden 1 y resolvemos en x_2:

$$\ddot{x}_2 + \omega_\circ^2 x_2 = -\frac{q}{m}\sum_{i=1}^{3}\frac{e^{\omega_i t}+e^{-\omega_i t}}{2} + bx_1^3 = -\frac{q}{m}\sum_{i=1}^{3}\frac{e^{\omega_i t}+e^{-\omega_i t}}{2} -$$

$$-\frac{bq^3}{8m^3}\sum_{ijk}\frac{E(\pm\omega_i)E(\pm\omega_j)E(\pm\omega_k)}{(\omega_\circ^2-\omega_i^2)(\omega_\circ^2-\omega_j^2)(\omega_\circ^2-\omega_k^2)}\frac{e^{\omega_l t}+e^{-\omega_l t}}{2} \Rightarrow$$

$$x_2(t) = -\frac{q}{m}\sum_{i=1}^{3}\frac{E(\omega_i)}{(\omega_\circ^2-\omega_i^2)}\frac{e^{\omega_i t}+e^{-\omega_i t}}{2} -$$

$$-\frac{bq^3}{8m^3}\sum_{ijk}\frac{E(\pm\omega_i)E(\pm\omega_j)E(\pm\omega_k)}{(\omega_\circ^2-\omega_i^2)(\omega_\circ^2-\omega_j^2)(\omega_\circ^2-\omega_k^2)(\omega_\circ^2-\omega_l^2)}\frac{e^{\omega_l t}+e^{-\omega_l t}}{2} \tag{16.9}$$

El primer término de esta última expresión no nos interesa ya que es el responsable de la polarización lineal. El segundo está asociado al mezclado de cuatro ondas. La polarización macroscópica asociada (considerando, de nuevo, que el número de átomos en resonancia es pequeño, por lo que $T_\circ \simeq \infty$) tiene la siguiente expresión:

$$P_F^{3)}(\omega_l) = -\frac{Nbq^4}{8dVm^3}\sum_{ijk}\frac{E(\pm\omega_i)E(\pm\omega_j)E(\pm\omega_k)}{(\omega_\circ^2-\omega_i^2)(\omega_\circ^2-\omega_j^2)(\omega_\circ^2-\omega_k^2)(\omega_\circ^2-\omega_l^2)}\frac{e^{\omega_l t}+e^{-\omega_l t}}{2} \tag{16.10}$$

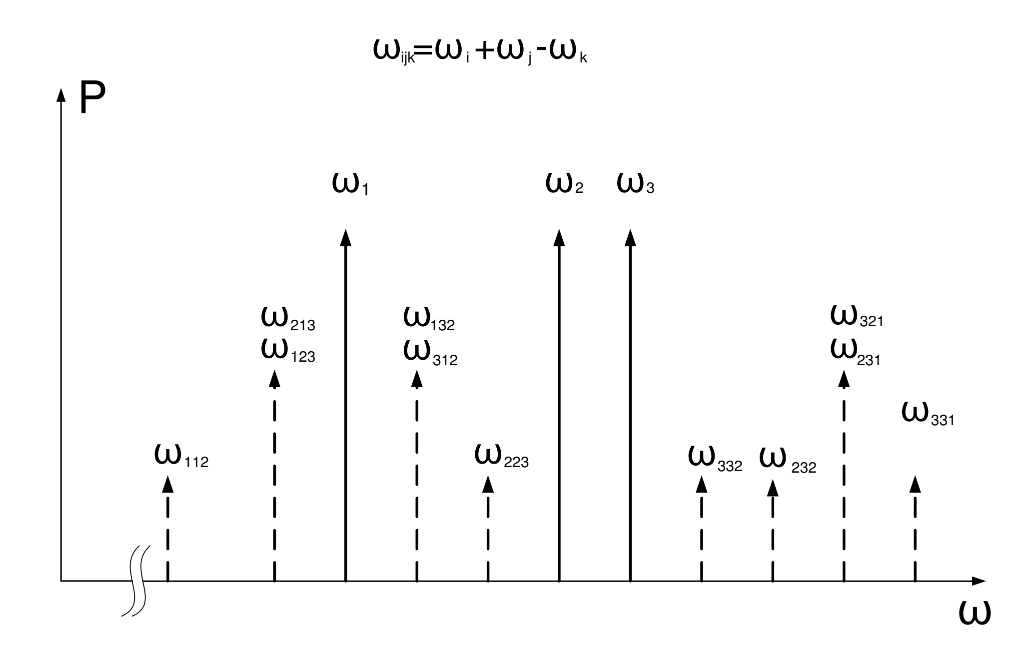


Figura 16.5: Señales espúreas producidas por el mezclado de cuatro ondas.

Evidentemente se tiene que cumplir $\omega_l = \pm\omega_i \pm \omega_j \pm \omega_k$. Como estamos estudiando cómo afecta el mezclado de cuatro ondas (FWM, del inglés "Four wave mixing") a los sistemas DWM, resulta claro que para que ω_l caiga dentro del ancho de banda de los canales transmitidos, se tiene que verificar que dos de los signos de (16.10) sean positivos y uno negativo. Como los canales están equiespaciado, ω_l tiene la misma pulsación que uno de ellos con lo que el FWM puede introducir unas diafonías entre canales muy perjudiciales. Además el número de señales que se producen por FWM es muy elevado, tal y como se muestra en la fig. {16.5}.

Supongamos que por la fibra se propagan tres canales de pulsaciones ω_1, ω_2 y ω_3. Una onda que se podría generar sería aquella cuya pulsación fuese $\omega_4 = \omega_1 + \omega_2 - \omega_3$. Sin embargo, para que esta onda se pueda propagar por la fibra es necesario que se verifique la siguiente condición:

$$\vec{k}_4 = \vec{k}_1 + \vec{k}_2 - \vec{k}_3 \tag{16.11}$$

Esta condición recibe el nombre de acuerdo de fase y su significado físico es que se tiene que conservar el momento total en este proceso de interacción entre los cuatro fotones [87]. La condición $\omega_4 = \omega_1 + \omega_2 - \omega_3$, a su vez, no representa más que el principio de conservación de la energía total [87]. La condición (16.11) se cumpliría exactamente en una fibra con dispersión cromática nula, ya que entonces la relación entre $\vec{k}$ y ω sería lineal y puesto que se cumple $\omega_4 = \omega_1 + \omega_2 - \omega_3$, automática se verificaría la relación (16.11). Puesto que el mezclado de cuatro onda puede introducir una diafonía muy importante, las fibras que se emplean en DWDM (las fibras de dispersión plana) tienen una dispersión residual de unos 5ps/nmKm, aún cuando podrían fabricarse con una dispersión mucho más baja. El motivo es impedir que se cumpla la condición de acuerdo de fase y reducir, así, el FWM [26].

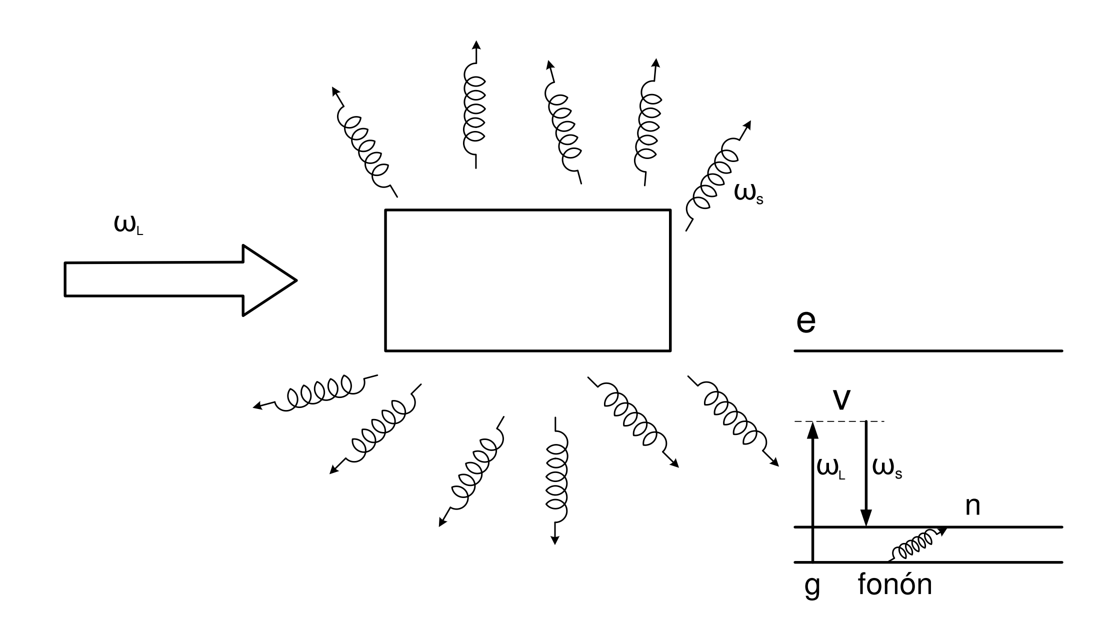


Figura 16.6: Scattering de fotones por efecto Raman espontáneo.

16.6. Efecto Raman estimulado (SRS)

Antes de estudiar el efecto Raman estimulado, vamos a describir el efecto Raman espontńeo [18]. Se trata de un proceso *no resonante*, lo que quiere decir que se absorbe un fotón que tiene un energía inferior a la necesaria para hacer pasar al átomo al estado excitado. El fotón absorbido y el electrón pasan a un estado virtual [3]. Cuando el electrón vuelve al estado fundamental éste, en algunos casos, tiene una energía ligeramente mayor que antes de producirse la transición, ya que ha absorbido un fonón óptico (los fonones son los cuantos asociados a las vibraciones de los enlaces del material y transportan tanto energía como cuasiimpulso). Por consiguiente, el fotón emitido tiene una energía ligeramente menor que la que tenía antes de producirse la transición al estado virtual, tal y como se muestra en la fig. {16.6}.

Si la luz está generada por un láser y es suficientemente potente se produce el denominado scattering Raman estimulado en el que un porcentaje importante de la luz se tranforma en la denominada onda Stokes (cuya pulsación es menor que la de la radiación incidente, $\omega_s < \omega$). Además, a diferencia de lo que ocurre con el scattering Raman espontáneo, la luz así producida no es isotrópica, sino que se propaga codireccionamente o contradireccionalmente con la radiación que excita el proceso. El motivo por el que se produce este fenómeno es que al propagarse por el medio tanto la intensa luz generada por el láser como la onda de Stokes, estos dos haces interaccionan con el medio material estimulando la creación de nuevos fonones ópticos, con lo que se produce una realimentación positiva [18]. El resultado es que una fracción importante de la luz láser se puede transformar en la onda de Stokes, tal como se muestra en la fig. {16.7}.

[3] Los estados virtuales no pueden medirse experimentalmente y su tiempo de existencia viene dado por la segunda relación de incertidumbre de Heissenberg: $\Delta E \Delta t \geq \hbar$.

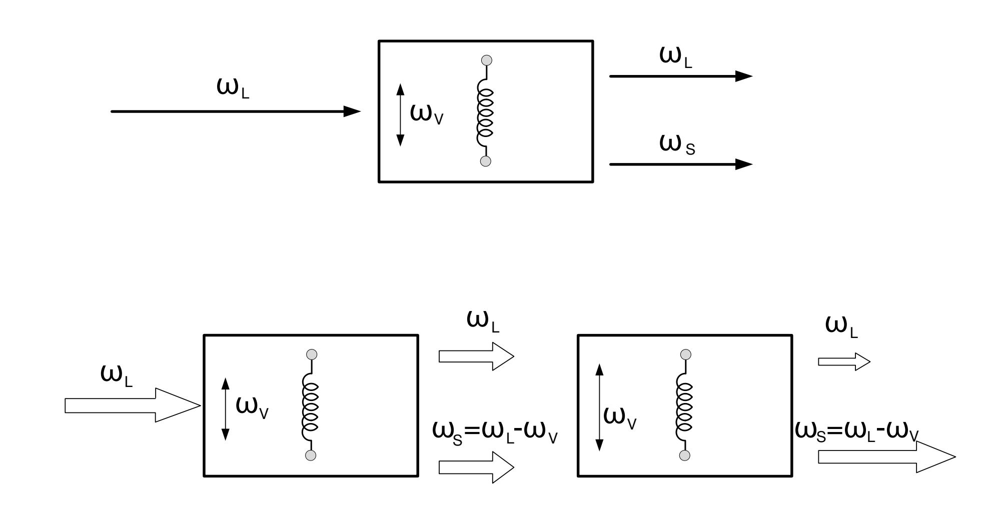


Figura 16.7: Realimentación positiva que da lugar al efecto Raman estimulado.

Al efecto Raman se le puede dar una interpretación clásica basada en el modelo de Lorentz del átomo. Este fenómeno ocurre porque los dipolos eléctricos que se inducen en el medio material depende ligeramente de la distancia interatómica. Por tanto, cuando los enlaces entre átomos vibran, el momento dipolar inducido en el medio material se modifica. El estado "n" excitado (por las vibraciones del enlace) no es, en realidad, un estado discreto, sino que forma una banda, de manera que se puede producir transferencia de energía desde un ω_L a todo un continuo de valores ω_s (el ancho de banda es del orden de una decena de THz). Aunque el efecto Raman estimulado se puede usar para implementar amplificadores distribuidos [5] inyectando una ω_L periódicamente a lo largo de la fibra, en los sistemas DWDM puede llegar a ser muy perjudicial. Esto es debido a que los canales de mayor frecuencia (menor longitud de onda) pueden actuar como potencia óptica que estimule el efecto Raman (con lo que se ven atenuados), mientras que los de menor frecuencia (mayor longitud de onda) son amplificados [26], tal y como se muestra en la fig. {16.8}. Evidentemente, se prduce SRS sólo cuando los pulsos de los distintos canales solapan. De nuevo, para minimizar el tiempo de solapamiento de los bits, se trabaja con fibras que tienen una dispersión cromática residual considerable y con un espaciado amplio entre canales. De todas maneras el efecto Raman limita la potencia óptica total que se puede transmitir por la fibra.

16.7. Efecto Brillouin estimulado (SBS)

El efecto Brillouin estimulado es parecido al SRS salvo que los estados vibratorios del sílice involucrados en este proceso son acústicos y no ópticos [18]. Por este motivo el ancho de banda de este fenómeno es muy estrecho y las distorsiones que provoca en los sistemas DWDM son siempre distorsiones *intracanal*. La forma de evitar el SBS es impedir que la potencia óptica en cada canal supere un determinado valor máximo [7].

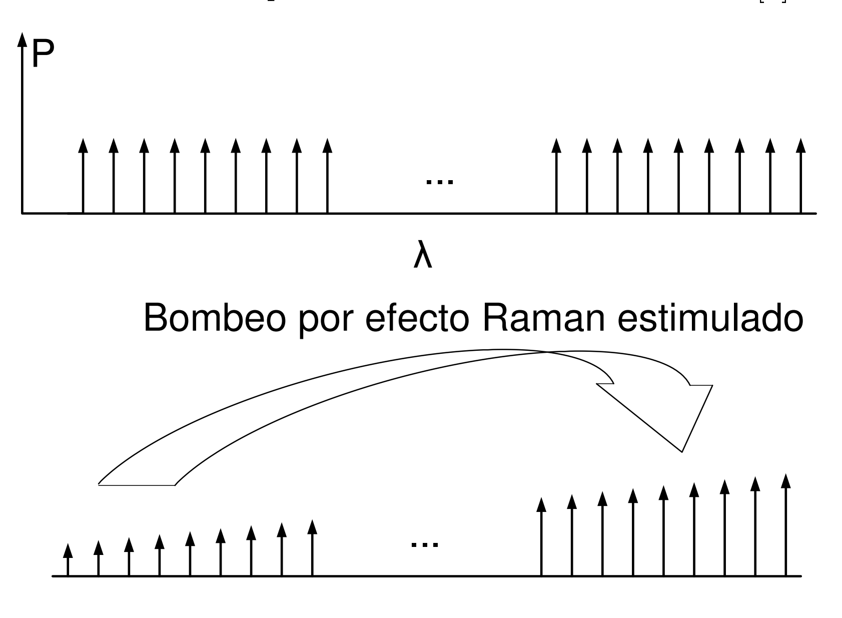


Figura 16.8: Degradación en un sistema DWDM por SRS.

Bibliografía

[1] S. Adachi, "Properties of Group-IV, III–V and II–VI Semiconductors", John Wiley and Sons, West Sussex (Reino Unido), 2005.

[2] R.B. Adler, A.C. Smith y R.L. Longini, "Introducción a la Física de los Semiconductores", Reverté, Barcelona (España), 1970.

[3] G.P. Agrawal, "Optical Communication Systems", John Wiley and Sons, Nueva York (EE.UU.), 1992.

[4] G.P. Agrawal, "Long-Wavelenth Semiconductor Lasers", Van Nostrand Reinhold, Nueva York (EE.UU.), 1986.

[5] G.P. Agrawal y C. Headley, "Raman Amplification in Fiber Optical Communication Systems", Elsevier Academic Press, San Diego (EE.UU.), 2005.

[6] G.P. Agrawal, "Nonlinear Fiber Optics", Academic Press, San Diego (EE.UU.), 1995.

[7] G.P. Agrawal, "Applications of Nonlinear Fiber Optics", Academic Press, San Diego (EE.UU.), 2001

[8] G.P. Agrawal, "Lightwave Technology: Components and Devices", John Wiley ans Sons, Nueva Jersey, (EE.UU.), 2004.

[9] A. Akhiezer y I. Akhiezer, "Électromagnétisme et ondes électomagnétiques", Editions Mir, Moscú (Ex-Unión Soviética), 1988.

[10] Z.I. Alferov y R.F. Kazarinov, "Semiconductor Laser with Electrical Pumping", U.R.S.S. Patente 181737, 1963.

[11] L. Allen y J.H. Eberly, "Optical Resonance and Two-level Atoms", Dover Publications, Nueva York (EE.UU.), 1987.

[12] N.W. Ashcroft y N.D. Mermin, "Solid State Physics", Saunders College, Philadelphia (EE.UU.), 1988.

[13] P.C. Becker, N.A. Olsson y J.R. Simpson, "Erbium-Doped Fiber Amplifiers, Fundamentals and Technology", Academic Press, San Diego (EE.UU.), 1999.

[14] F. Bertein, "Bases de l'Électronique Quantique Tome 1", Éditions Eyrolles, Paris (Francia), 1965.

[15] F. Bertein, "Bases de l'Électronique Quantique Tome 2", Éditions Eyrolles, Paris (Francia), 1965.

[16] A. Bjarklev, "Optical Fiber Amplifiers: Design and System Applications", Artech House, Norwood (EE.UU.), 1993.

[17] M. Born y E. Wolf, "Principles of Optics", Pergamon Press, Exeter (Reino Unido), 1986.

[18] R.W. Boyd, "Nonlinear Optics", Academis Press, San Diego (EE.UU.), 2003.

[19] P.K. Cheo, "Fiber Optics and Optoelectronics", Prentice Hall, Nueva Jersey (EE.UU.), 1990.

[20] W.W. Chow y S.W. Koch, "Semicondutor-Laser Fundamentals: Physics of the Gain Material", Springer, Heidelberg (Alemania), 1999.

[21] S.H. Chuang, "Physics of Optoelectronic Devices", John Wiley and Sons, Nueva York (EE.UU.), 1995.

[22] C. Cohen-Tannoudji, J. Dupont-Roc y G. Grynberg, "Photons and Atoms—Introduction to Quantum Electrodynamics", John Wiley and Sons, Nueva York (EE.UU.), 1997.

[23] C. Cohen-Tannoudji, J. Dupont-Roc y G. Grynberg, "Atom–Photon Interactions: Basic Processes and Interactions", John Wiley and Sons, Nueva York (EE.UU.), 1997.

[24] L.A. Coldren y S.W. Corzine, "Diode Lasers and Photonic Integrated Circuits", John Wiley and Sons, Nueva York (EE.UU.), 1995.

[25] J.M. Connelly, "Semiconductor Optical Amplifiers", Kluwer Academic Press, Nueva York (EE.UU.), 2002.

[26] M. Cvijetic, "Optical Transmission: Systems Engineering", Artech House, Norwood (EE.UU.), 2004.

[27] E. Desurvire, "Erbium-Doped Fiber Amplifiers", John Wiley and Sons, Nueva York (EE.UU.), 1994.

[28] P.A.M. Dirac, "The Quantum Theory of the Emission and Absorption of Radiation", Proceedings of the Royal Society of London, Series A, vol.114, 1927.

[29] T. Erdogan, "Fiber Grating Spectra", Journal of Lightwave Technology, vol.15, n. 8, Agosto 1997.

[30] W. van Etten y J. van der Plaats, "Fundamentals of Optical Communications", Prentice Hall, Cambridge (Reino Unido), 1991.

[31] R. Feced, M.N. Zervas y M.A. Muriel, "An Efficient Scattering Algorithm for the Design of Nonuniform Fiber Bragg Gratings", IEEE Journal of Quantum Electronics, vol. 35, n. 8, Agosto 1999.

[32] R.P. Feynman, "Quantum Electrodynamics", Addison Wesley, (EE.UU.), 1961.

[33] E. Garmire y A. Kost (editores), "Nonlinear Optics in Semiconductors I", Academic Press, San Diego (EE.UU.), 1999.

[34] H. Ghafouri–Shiraz, "Distributed Feedback Laser Diodes and Optical Tunable Lasers, John Wiley and Sons, West Sussex (Reino Unido), 2003.

[35] T. Georges, "Solitons dans les Fibres Optiques", Techniques de l'Ingenieur, traité Électronique, vo. 1 n. E1985, 1999.

[36] C.R. Giles y E. Desurvire, "Modeling Erbium-Doped Amplifiers", Journal of Lightwave Technology, vol. 9, n. 2, Febrero 1991.

[37] M. Gilmore, "Fibre Optic Cabling", Newness, Oxford (Reino Unido), 1996.

[38] D. Glodge, "Weakly Guiding Fibers", Applied Optics, vol. 10, pp. 2252–2258, 1971.

[39] K.O. Hill, Y. Fujii, D.C. Johnson y B. S. Kawasaki, "Photosensitivity in optical fiber waveguides: Application to reflection filter fabrication", Applied Physics Letters, vol. 32, n. 10, Mayo 1978

[40] J.P. Gordon y H. Kogelnik, "PMD Fundamentals: Polarization Mode Dispersion in Optical Fibers", Proceedings of the National Academy of Sciences, vol. 97, n. 8, Abril 2000.

[41] J. Gowar, "Optical Communication Systems", Prentice Hall, Exeter (Reino Unido), 1984.

[42] H. Haken, "Light: Waves, Photons, Atoms", North Holland Publishing Company, Amsterdam (Paises Bajos), 1981.

[43] H. Haken, "Light: Laser Light Dynamics", North Holland Publishing Company, Amsterdam (Paises Bajos), 1984.

[44] P. Harrison, "Quantum Wells, Wires and Dots", John Wiley and Sons, West Sussex (Reino Unido), 2005.

[45] A. Hasegawa y Y. Kodama, "Solitons in Optical Communications", Clarendon Press, Oxford (Reino Unido), 1995.

[46] A. Hasegawa y M. Matsumoto, "Optical Solitons in Fibers", Springer-Verlag, Berlin (Alemania), 2003.

[47] H. A. Haus, "Optical Fiber Solitons, Their Properties and Uses", Proceedings if the IEEE, vol. 81, n. 7, Julio 1993.

[48] W.H. Hayward, "Introduction to Radio Frequency Design", ARRL, (EE.UU.), 1996.

[49] E. Hecht "Óptica", Addison Wesley, Madrid (España), 2000.

[50] W. Heitler, "The Quantum Theory of Radiation", Dover Publications, Nueva York (EE.UU.), 1984.

[51] C.H. Henry, "Theory of the Linewidth of Semiconductor Lasers", IEEE Journal of Quantum Electronics, vol. QE-18, n. 2. Febrero 1982.

[52] P.C.D. Hobbs, "Building Electrooptical Systems (Making it all Work)", John Wiley and Sons, Nueva York (EE.UU.), 2000.

[53] S. Huard, "Polarization de la Lumière", Masson, París (Francia), 1993.

[54] E. Iannone, F. Matera, A. Mecozzi y M. Settembre, "Nonlinear Optical Communication Networks", John Wiley and Sons, Nueva York (EE.UU.), 1998.

[55] K. Iizuka, "Elements of Photonics I: In Free Space and Special media", John Wiley and Sons, Nueva York (EE.UU.), 2002.

[56] K. Iizuka, "Elements of Photonics II: For Fiber and Integrated Optics", John Wiley and Sons, Nueva York (EE.UU.), 2002.

[57] G. Jacobsen, "Noise in Digital Optical Transmission Systems", Artech House, Boston (EE.UU.), 1994.

[58] I.P. Kaminov y T. Li (editores), "Optical Fiber Telecommunications IVA: Components", Academic Press, San Diego (EE.UU.), 2002.

[59] I.P. Kaminov, T. Li y A.E. Willner (editores), "Optical Fiber Telecommunications V: Components and Subsystems", Academic Press, San Diego (EE.UU.), 2008.

[60] R. Kashyap, "Fiber Bragg Gratings", Academic Press, San Diego (EE.UU.), 1999.

[61] M.V. Klibanov, P.E. Sacks y A.V. Tikhonravov, "The Phase Retrieval Problem", Inverse problems 11, Institute of Physics Publishing, 1995.

[62] C. Klingshirn, "Semiconductor Optics", Springer Verlag, Berlín, Alemania, 2005.

[63] H. Kroemer. "A Proposed Class of Heterojunction Lasers", Proc. IEEE vol. 51, 1963.

[64] M. Levinshtein, J. Kostamovaara y S. Vainshtein, "Breakdown Phenomena in Semiconductors and Semiconductor Devices", World Scientific, Singapur (Singapur), 2005.

[65] G. Lifante, "Integrated Photonics: Fundamentals", John Wiley and Sons, West Sussex (Reino Unido), 2003.

[66] M.M.K. Liu, "Principles and Applications of Optical Communications", Irwin, (EE.UU.), 1996.

[67] R. Loudon, "The Quantum Theory of Light", Oxford University Press, Nueva York (EE.UU.), 2000.

[68] H.A. Macleod, "Thin–Film Optical Filters", Institute of Physics Publishing, Bristol (Reino Unido), 2001.

[69] W. Mandel y E. Wolf, "Optical Coherence and Quantum Optics", Cambridge University Press, Cambridge (Reino Unido), 1995.

[70] D. Marcuse, "Theory of Dielectric Optical Waveguides", Acadamic Prress, Nueva York (EE.UU.), 1974.

[71] J.P. McKelvey, "Física del Estado Sólido y Semiconductores", Limusa, México D.F. (México), 1991.

[72] D.J.G. Mestdagh, "Fundamentals of Multiaccess Optical Fiber Networks", Artech House, Boston (EE.UU.), 1995.

[73] P.W. Milonni, "The Quantum Vacuum", Academic Press, Londres (Reino Unido), 1994.

[74] T. Numai, "Fundamentals of Semiconductor Lasers", Springer Verlag, Nueva York (EE.UU.), 2004.

[75] K. Okamoto, "Fundamentals of Optical Waveguides", Academic Press, Nueva York (EE.UU.), 2005.

[76] T. Okoshi, "Optical Fibers", Academic Press, San Diego (EE.UU.), 1982.

[77] R. van Overstraeten y P. Heremans, "Semiconductor Devices", Uitgeverij Acco, Lovaina (Bélgica), 1998.

[78] F.L. Pedrotti y L.S. Pedrotti, "Introduction to Optics", Prentice-Hall, Nueva Yersey (EE,UU.), 1987.

[79] L. Poladian, "Simple Grating Synthesis Algorithm", Optics Letters, vol. 25, n. 11, Junio 2000.

[80] C.D. Poole, "Statistical Treatment of Polarization Dispersion in Single-Mode Fiber", Optics Letters, vol. 13, n. 8. Agosto 1998.

[81] C.D. Poole, "Measurement of Polarization-Mode Dispersion in Single-Mode Fibers with Random Mode Coupling", Optics Letters, vol. 14, n. 10. Mayo 1989.

[82] C.D. Poole, J.H. Winters y J.A. Nagel, "Dynamical Equation for Polarization Dispersion", Optics Letters, vol. 16, n. 6. Marzo 1991.

[83] J.G. Proakis, "Digital Communications", MgGraw-Hill, Nueva York (EE.UU.), 1983.

[84] R. Ramaswami y K.N. Sivarajan, "Optical Networks: A Practical Perspective", Morgan Kaufmann Publishers, San Francisco (EE.UU.), 1998.

[85] D.W. Ricker, "Echo Signal Processing", Kluwer Academic Publisher, Nueva York (EE.UU.), 2003.

[86] R. Sabella y P. Lugli, "High Speed Optical Communications", Kluwer Academic Publisher, Dordrecht (Paises Bajos), 2002.

[87] B.E. Saleh y M.C. Teich, "Fundamentals of Photonics", John Wiley and Sons, Nueva York (EE.UU.), 1991.

[88] D. Schwarzenbach, "Cristallography", John Willey and Sons, West Sussex (Reino Unido), 1996.

[89] M.O. Scully y M.S. Zubairy, "Quantum Optics", Cambridge University Press, Cambridge (Reino Unido), 2002.

[90] K.V. Shalimova, "Física de los Semiconductores", Editorial MIR, (U.R.S.S.), 1975.

[91] A.L. Schawlow y C.H. Townes, "Infrared and Optical Masers", Physical Review, vol. 112, n. 6. Diciembre 1958.

[92] A.E. Siegman, "Lasers", University Science Books, California (EE.UU.), 1986.

[93] J. Singh, "Electronic and Optoelectronic Properties of Semiconductor Structures", Cambridge University Press, Cambridge (Reino Unido), 2003.

[94] J. Skaar, "Synthesis of Fiber Bragg Gratings for Use in Transmission", Journal of the Optical Society of America A, vol. 18, n. 3, Marzo 2001.

[95] J. Skaar, L. Wang y T. Erdogan, "On the Synthesis of Fiber Gratings by Layer Peeling", IEEE Journal of Quantum Electronics, vol. 37, n. 2, Febrero 2001.

[96] H. Uetsuka, "AWG Technologies for Dense WDM Applications", IEEE Journal of Selected Topics in Quantum Electronics, vol. 10, n. 2, Marzo–Abril 2004.

[97] V.M. Ustinov, A.E. Zhukov, A.Y. Egorov y M.A. Maleev, "Quantum Dot Lasers", Oxford University Press, Nueva York (EE.UU.), 2003.

[98] J.T. Verdeyen, "Laser Electronics", Prentice Hall, Nueva Jersey (EE.UU.), 1989.

[99] Z. Wang y S. Fan, "Optical Circulators in Two-Dimensional Magneto-Optical Photonic Crystals", Optics Letters, vol. 50, n. 15, Agosto 2005.

[100] G.B. Whitham, "Linear and Nonlinear Waves", John Wiley and Sons, Nueva York, (EE.UU.), 1974.

[101] G. Willems, "Fundamentals of Solid State Physics for Semiconductors", Uitgeverij Acco, Lovaina (Bélgica), 1995.

[102] D. Wood, "Optoelectronic Semiconductor Devices", Prentice Hall, Hertfordshire (Reino Unido), 1994.

[103] A. Yariv, "Coupled-Mode Theory for Guided-Wave Optics", IEEE Journal of Quantum Electronics, vol. QE-9, n. 9, Septiembre 1973.

[104] A. Yariv, "Optical Electronics in Modern Communications", Oxford University Press, Nueva York (EE.UU.), 1997.

[105] A. Yariv, "Quantum Electronics", John Wiley and Sons, Nueva York (EE.UU.), 1989.

[106] A. Yariv y P. Yeh, "Optical Waves in Cristals", John Wiley ans Sons, Nueva York (EE.UU.), 1984.

[107] P. Yeh, "Optical Waves in Layer Media", John Wiley and Sons, Nueva York (EE.UU.), 1988.

[108] V.E. Zakharov y A.B. Shabat, "Exact Theory of Two-Dimensional Self Focusing and One-Dimensional Modulation of Waves in Nonlinear Media", Soviet Physic Letters, vol. 34, pp. 62–69, 1972.

[109] P.S. Zory, "Quantum Well Lasers", Academic Press, San Diego (EE.UU.), 1993.